Top 10 Words Per Item (index-based labels)

=== item_0 ===

sec: 5, proposal: 4, new: 2, securities: 2, exchange: 2, like: 2, wall: 2, street: 2, industry: 2, understand: 2

=== item_1 ===

federal: 4, reserve: 4, financial: 4, said: 4, crisis: 3, senator: 3, banking: 3, committee: 3, criticised: 2, enough: 2

=== item_2 ===

internet: 6, service: 6, court: 4, ban: 4, leesburg: 3, first: 3, stuck: 2, traffic: 2, pike: 2, northern: 2

=== item_3 ===

tax: 6, code: 5, coalition: 4, corporate: 4, said: 4, boustany: 4, overhaul: 3, microsoft: 2, chamber: 2, business: 2

=== item_4 ===

government: 6, said: 6, blouin: 5, new: 2, car: 2, firm: 2, cars: 2, consumer: 2, choice: 2, department: 2

=== item_5 ===

bill: 5, would: 4, debate: 3, iran: 3, authorize: 2, use: 2, force: 2, nuclear: 2, opponents: 2, war: 2

=== item_6 ===

said: 5, yellen: 4, federal: 3, cyberattacks: 3, careful: 3, banks: 2, would: 2, think: 2, reserve: 2, one: 2

=== item_7 ===

rules: 8, export: 6, encryption: 5, commerce: 5, department: 5, products: 4, congress: 2, used: 2, protect: 2, customer: 2

=== item_8 ===

government: 6, letter: 5, said: 4, encryption: 3, commerce: 3, protect: 3, privacy: 3, congress: 2, department: 2, americans: 2

=== item_9 ===

federal: 4, bank: 4, fed: 4, regulators: 3, new: 3, technology: 3, banks: 3, financial: 2, use: 2, fraud: 2

=== item_10 ===

bank: 3, markets: 2, ruble: 2, said: 2, einhorn: 2, economy: 2, russian: 2, currency: 2, central: 2, qe: 2

=== item_11 ===

index: 5, said: 4, government: 4, gain: 3, ftse: 3, percent: 3, stock: 2, thursday: 2, czech: 2, since: 2

=== item_12 ===

percent: 4, points: 3, investors: 2, dow: 2, jones: 2, p: 2, nasdaq: 2, composite: 2, think: 2, panic: 2

=== item_13 ===

dollar: 4, stock: 3, index: 3, czech: 3, exchange: 2, gain: 2, republic: 2, percent: 2, major: 2, market: 2

=== item_14 ===

czech: 5, consumer: 5, november: 5, percent: 5, inflation: 4, prices: 3, republic: 3, survey: 2, central: 2, bank: 2

=== item_15 ===

bill: 6, chamber: 5, czech: 3, parliament: 3, parties: 3, would: 3, lower: 3, house: 3, deputies: 3, klaus: 2

=== item_16 ===

said: 5, iran: 4, klaus: 3, agreement: 3, minister: 2, deal: 2, countries: 2, nuclear: 2, iranian: 2, zarif: 2

=== item_17 ===

czech: 5, percent: 4, gas: 4, republic: 3, economy: 3, growth: 2, expected: 2, next: 2, years: 2, likely: 2

=== item_18 ===

going: 3, central: 3, bank: 3, banks: 3, ecb: 3, said: 2, gros: 2, announce: 2, monday: 2, policy: 2

=== item_19 ===

czech: 5, republic: 5, country: 4, may: 3, entry: 3, eurovision: 2, roma: 2, contest: 2, ljubljana: 2, city: 2

=== item_20 ===

property: 4, report: 4, said: 4, commercial: 3, market: 3, export: 3, ambitions: 3, uk: 3, industry: 3, world: 3

=== item_21 ===

ritchie: 6, manager: 5, firm: 5, grenfell: 4, former: 4, deutsche: 3, morgan: 3, fund: 3, money: 3, nicola: 2

=== item_22 ===

ecb: 5, next: 5, cap: 4, expected: 3, per: 3, cent: 3, gdp: 3, week: 3, foreign: 2, exchange: 2

=== item_23 ===

fund: 7, superwoman: 4, horlick: 3, director: 2, honored: 2, years: 2, cause: 2, created: 2, died: 2, manager: 1

=== item_24 ===

fund: 6, superwoman: 4, horlick: 4, said: 4, made: 3, pension: 2, decision: 2, exhaustive: 2, search: 2, applicants: 2

=== item_25 ===

grenfell: 4, trial: 4, told: 4, action: 3, group: 3, former: 2, appeared: 2, stand: 2, court: 2, years: 2

=== item_26 ===

grenfell: 6, tower: 5, bank: 4, new: 3, said: 3, deutsche: 2, unlikely: 2, offer: 2, nicola: 2, horlick: 2

=== item_27 ===

london: 4, bank: 3, offices: 3, banks: 3, sfo: 3, deutsche: 2, morgan: 2, fraud: 2, settlement: 2, includes: 2

=== item_28 ===

bank: 2, germany: 1, dresdner: 1, dutch: 1, abn: 1, amro: 1, holding: 1, nv: 1, also: 1, part: 1

=== item_29 ===

state: 5, investment: 4, person: 4, would: 4, banking: 3, billion: 3, deal: 3, said: 3, giant: 2, merrill: 2

=== item_30 ===

court: 6, liu: 3, appointment: 3, party: 3, zhou: 3, former: 3, chinese: 2, filed: 2, lawsuit: 2, communist: 2

=== item_31 ===

china: 6, party: 4, hong: 3, kong: 3, communist: 3, said: 3, deng: 3, hailed: 2, selection: 2, mainland: 2

=== item_32 ===

united: 4, states: 4, nuclear: 4, china: 3, said: 3, weapons: 3, party: 2, jiang: 2, give: 2, since: 2

=== item_33 ===

internet: 7, said: 5, government: 5, foreign: 4, firms: 4, open: 3, insurance: 3, investment: 3, ministry: 3, commerce: 3

=== item_34 ===

china: 3, composite: 3, top: 3, said: 2, overheating: 2, sharp: 2, rise: 2, chinese: 2, shares: 2, investors: 2

=== item_35 ===

said: 4, companies: 4, rules: 4, china: 3, new: 3, products: 2, list: 2, names: 2, symbols: 2, use: 2

=== item_36 ===

rates: 4, fed: 4, would: 4, said: 3, interest: 3, economy: 3, china: 2, chinese: 2, warned: 2, inflation: 2

=== item_37 ===

foreign: 7, china: 6, government: 4, markets: 3, companies: 3, direct: 3, investment: 3, said: 3, insurance: 2, stock: 2

=== item_38 ===

wang: 7, chinese: 3, donghai: 3, china: 3, dissident: 2, trying: 2, political: 2, beijing: 2, said: 2, arrest: 2

=== item_39 ===

china: 9, rights: 7, human: 6, united: 5, states: 5, said: 5, groups: 4, campaign: 3, supporting: 3, oppose: 3

=== item_40 ===

news: 7, journal: 5, corp: 3, murdoch: 3, company: 3, wall: 3, street: 3, media: 2, group: 2, new: 2

=== item_41 ===

home: 8, loan: 7, business: 7, bank: 6, banking: 4, growth: 3, cba: 3, broader: 3, sector: 3, per: 3

=== item_42 ===

housing: 5, high: 4, australia: 3, bubbles: 3, said: 3, banking: 3, financial: 3, sector: 3, cba: 2, regulation: 2

=== item_43 ===

economy: 5, quarter: 5, percent: 4, said: 3, pace: 3, second: 2, first: 2, reuters: 2, poll: 2, survey: 2

=== item_44 ===

year: 6, bank: 5, quarter: 5, billion: 5, significant: 3, cba: 2, strong: 2, second: 2, positive: 2, outlook: 2

=== item_45 ===

decision: 4, group: 3, said: 3, daily: 3, rural: 2, press: 2, publications: 2, would: 2, sophisticated: 2, new: 2

=== item_46 ===

said: 5, bank: 4, policy: 4, fed: 4, murray: 3, rate: 3, central: 2, stance: 2, positive: 2, interest: 2

=== item_47 ===

deal: 5, network: 4, company: 4, would: 4, australian: 3, seven: 3, major: 3, television: 2, billion: 2, pressure: 2

=== item_48 ===

company: 5, said: 5, brambles: 3, accc: 3, act: 3, pallet: 2, shipment: 2, practices: 2, believes: 2, hire: 1

=== item_49 ===

olympic: 7, project: 6, government: 5, said: 5, bid: 4, billion: 3, loan: 3, federal: 3, australian: 3, would: 3

=== item_50 ===

looking: 6, said: 5, company: 4, sell: 4, businesses: 4, still: 3, buyer: 3, transition: 2, independent: 2, division: 2

=== item_51 ===

said: 6, buy: 4, financial: 4, deal: 3, person: 3, company: 3, transaction: 3, finance: 2, wall: 2, street: 2

=== item_52 ===

mercury: 4, stock: 4, million: 3, bank: 3, friday: 2, company: 2, said: 2, investors: 2, percent: 2, trading: 2

=== item_53 ===

said: 4, fbi: 3, mercury: 3, san: 3, francisco: 3, company: 3, monday: 2, statement: 2, field: 2, office: 2

=== item_54 ===

tax: 8, said: 5, new: 4, thornton: 4, rate: 3, company: 2, multinationals: 2, renaissance: 2, companies: 2, profits: 2

=== item_55 ===

percent: 4, change: 3, market: 3, stock: 3, strong: 2, said: 2, dow: 2, jones: 2, industrial: 2, average: 2

=== item_56 ===

internet: 6, said: 4, khosla: 3, government: 3, decide: 3, inc: 2, also: 2, given: 2, regulate: 2, communications: 2

=== item_57 ===

said: 5, mercury: 3, employees: 3, thursday: 2, overtime: 2, cut: 2, difficult: 2, company: 2, finance: 1, suspending: 1

=== item_58 ===

bank: 8, mercury: 3, said: 3, company: 3, dispute: 2, lawsuit: 2, spokesman: 2, requirement: 2, finance: 1, thursday: 1

=== item_59 ===

company: 8, said: 7, bank: 3, statement: 3, request: 3, major: 2, stores: 2, time: 2, matter: 2, reliance: 1

=== item_60 ===

said: 6, companies: 5, market: 3, offer: 3, number: 3, merger: 2, combined: 2, value: 2, billion: 2, potential: 2

=== item_61 ===

police: 6, bissonnette: 5, said: 4, thursday: 3, two: 3, alexandre: 3, father: 3, expected: 2, charges: 2, disappearance: 2

=== item_62 ===

power: 7, said: 7, outage: 3, system: 3, supply: 2, happen: 2, macpherson: 2, canadian: 2, association: 2, company: 2

=== item_63 ===

city: 5, businesses: 4, toronto: 3, new: 3, welcoming: 2, law: 2, old: 2, requires: 2, mayor: 2, tory: 2

=== item_64 ===

city: 6, ford: 6, canada: 3, toronto: 3, shakeup: 2, rob: 2, remembered: 2, things: 2, office: 2, move: 2

=== item_65 ===

decision: 5, united: 4, states: 4, wto: 4, membership: 3, indonesia: 3, said: 2, country: 2, made: 2, seek: 2

=== item_66 ===

company: 8, meeting: 4, molson: 3, annual: 3, april: 2, new: 2, york: 2, website: 2, report: 2, would: 2

=== item_67 ===

employee: 5, restaurant: 4, video: 4, police: 3, said: 3, wednesday: 3, man: 3, hospital: 3, bag: 3, violent: 2

=== item_68 ===

canada: 5, public: 4, canadian: 4, strippers: 3, people: 3, dance: 2, naked: 2, long: 2, permit: 2, broadcasting: 2

=== item_69 ===

de: 6, guzman: 6, park: 5, death: 4, family: 4, national: 3, service: 3, said: 3, injuries: 2, trail: 2

=== item_70 ===

safety: 6, rule: 4, cars: 4, motor: 3, used: 3, said: 3, administration: 3, certified: 2, vehicles: 2, director: 2

=== item_71 ===

union: 6, workers: 4, chrysler: 4, said: 4, auto: 3, likely: 3, aam: 3, canadian: 2, boost: 2, power: 2

=== item_72 ===

said: 4, quarter: 4, gm: 4, expected: 3, sell: 3, million: 2, net: 2, profit: 2, fiscal: 2, first: 2

=== item_73 ===

police: 6, said: 5, dealership: 3, greene: 3, shot: 3, man: 2, shooting: 2, ford: 2, louis: 2, officers: 2

=== item_74 ===

company: 6, brands: 5, breakfast: 4, said: 4, food: 3, yum: 3, companies: 3, lunch: 2, combined: 2, stronger: 2

=== item_75 ===

kellogg: 7, said: 6, would: 6, breakfast: 5, offerings: 4, broaden: 3, new: 3, products: 3, companies: 3, food: 2

=== item_76 ===

workers: 6, carolina: 4, said: 4, agreement: 3, covers: 3, south: 3, ratified: 2, contract: 2, first: 2, american: 2

=== item_77 ===

discount: 6, offer: 4, food: 3, services: 3, said: 3, companies: 2, lure: 2, customers: 2, discounts: 2, even: 2

=== item_78 ===

retail: 5, sales: 3, percent: 3, according: 3, industry: 3, sears: 2, kmart: 2, boosted: 2, said: 2, gain: 2

=== item_79 ===

company: 6, billion: 4, chrysler: 3, plant: 3, move: 3, said: 3, transaction: 2, new: 2, marchionne: 2, allow: 2

=== item_80 ===

lucasvarity: 4, shares: 4, company: 3, deal: 3, would: 3, market: 3, tokyo: 3, parts: 3, companies: 2, japanese: 2

=== item_81 ===

sales: 5, landline: 4, said: 4, market: 4, phones: 3, televisions: 3, set: 3, next: 2, uk: 2, rise: 2

=== item_82 ===

court: 4, company: 4, said: 4, us: 4, equipment: 4, group: 3, block: 3, wireless: 3, communications: 2, blackwater: 2

=== item_83 ===

said: 5, racal: 4, united: 4, states: 4, antitrust: 4, million: 2, plant: 2, transaction: 2, car: 2, european: 2

=== item_84 ===

company: 7, said: 5, yen: 4, billion: 3, profit: 3, matsuyama: 3, hanson: 2, increase: 2, percent: 2, year: 2

=== item_85 ===

billion: 8, company: 5, quarter: 4, debt: 4, said: 4, rise: 3, net: 3, pounds: 3, percent: 3, profit: 2

=== item_86 ===

european: 5, uk: 3, government: 3, union: 3, britain: 2, eu: 2, amendment: 2, article: 2, vote: 2, british: 2

=== item_87 ===

government: 4, would: 4, said: 4, criminal: 3, israeli: 3, british: 2, friday: 2, refer: 2, international: 2, court: 2

=== item_88 ===

marijuana: 5, companies: 4, government: 3, would: 3, letter: 3, refer: 2, issue: 2, legalization: 2, world: 2, trade: 2

=== item_89 ===

eu: 6, free: 4, friday: 3, movement: 3, trade: 3, said: 3, british: 2, government: 2, would: 2, uk: 2

=== item_90 ===

year: 7, billion: 7, full: 5, vlsi: 4, percent: 4, expects: 3, quarter: 3, revenue: 3, said: 3, fourth: 2

=== item_91 ===

said: 7, company: 6, percent: 4, nelkes: 4, stock: 3, washington: 3, friday: 2, would: 2, workforce: 2, shutter: 2

=== item_92 ===

administration: 4, based: 3, information: 3, ap: 3, said: 3, white: 3, house: 3, reporting: 3, china: 3, story: 2

=== item_93 ===

percent: 7, shares: 4, trading: 3, analysts: 3, year: 3, company: 3, profit: 3, billion: 3, intel: 3, value: 2

=== item_94 ===

said: 6, access: 4, computer: 3, gained: 3, email: 3, phishing: 3, intrepid: 3, network: 2, hackers: 2, customer: 2

=== item_95 ===

company: 7, computer: 5, associates: 4, sell: 3, ammunition: 3, said: 3, would: 3, decision: 2, allow: 2, guns: 2

=== item_96 ===

company: 5, said: 4, cable: 4, cascade: 3, new: 3, million: 3, revenue: 2, quarter: 2, customers: 2, doubling: 2

=== item_97 ===

billion: 6, said: 4, company: 3, net: 3, million: 3, revenue: 3, shares: 2, trading: 2, quarterly: 2, loss: 2

=== item_98 ===

company: 4, said: 4, year: 4, spending: 3, percent: 3, consumer: 2, office: 2, equipment: 2, revenue: 2, earlier: 2

=== item_99 ===

percent: 4, fed: 3, stocks: 2, economic: 2, p: 2, last: 2, month: 2, investors: 2, rose: 2, nasdaq: 2

Stylistic Analysis

Total unique words: 3127

Total word count: 19078

Type Token Ratio (TTR): 16.39%

Total stop-words count: 7046

Most common Parts of Speech n-grams:

1. ('DT', 'NN'), 1293

2. ('NN', 'IN'), 942

3. ('IN', 'DT'), 830

4. ('JJ', 'NN'), 825

5. ('NN', 'NN'), 681

6. ('DT', 'JJ'), 572

7. ('NN', ','), 534

8. ('IN', 'NN'), 415

9. ('NN', 'VBD'), 387

10. ('TO', 'VB'), 381

11. ('NN', '.'), 365
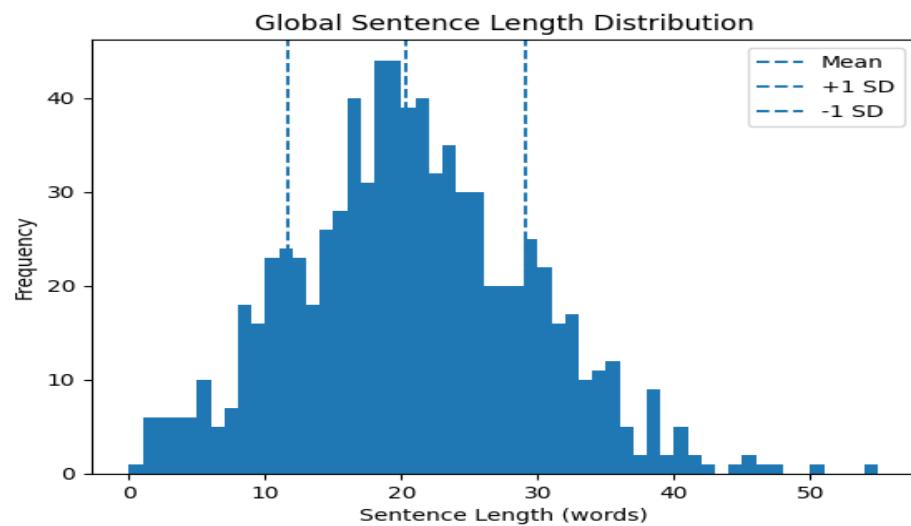
12. ('.', 'DT'), 287

13. ('JJ', 'NNS'), 276
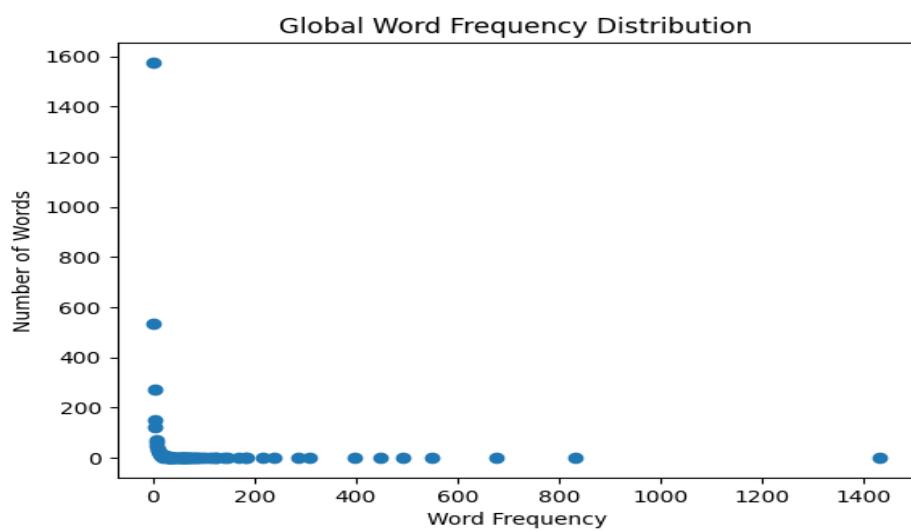
14. ('NNS', 'IN'), 232

15. ('IN', 'JJ'), 209
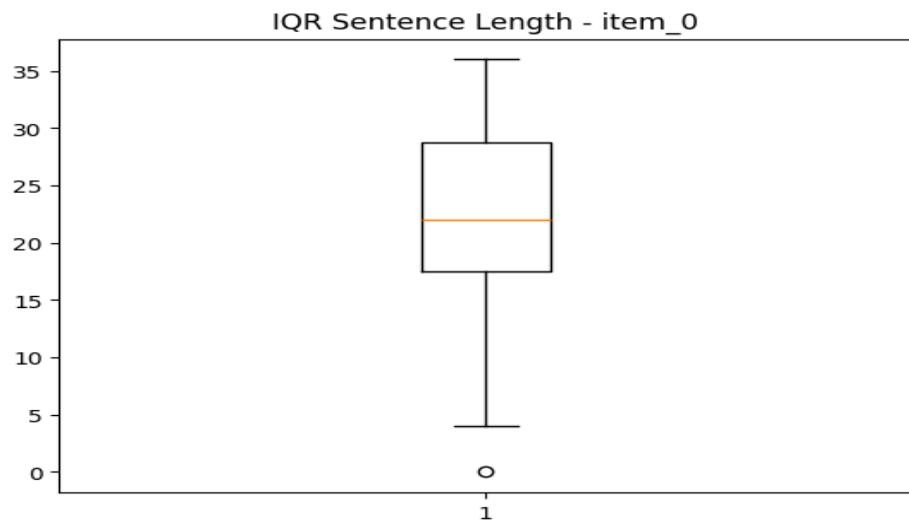
global_sentence_length_distribution.png
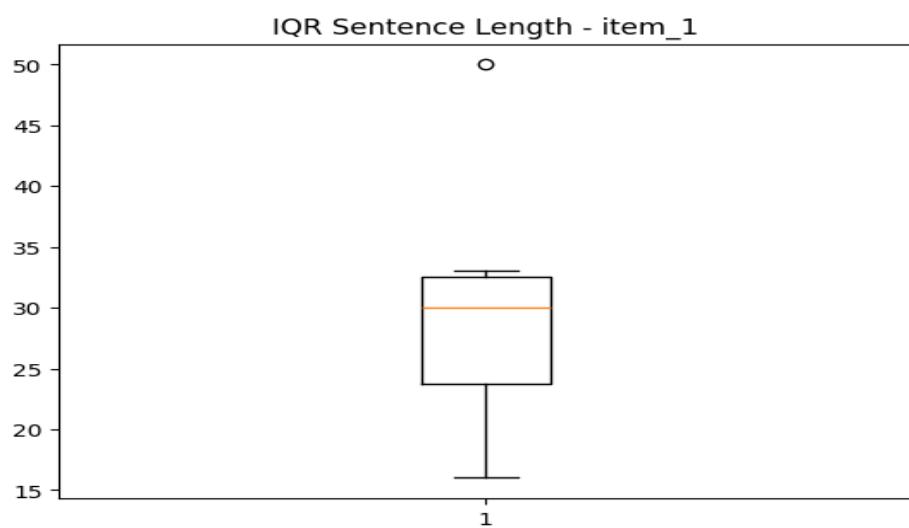
global_word_freq_distribution.png
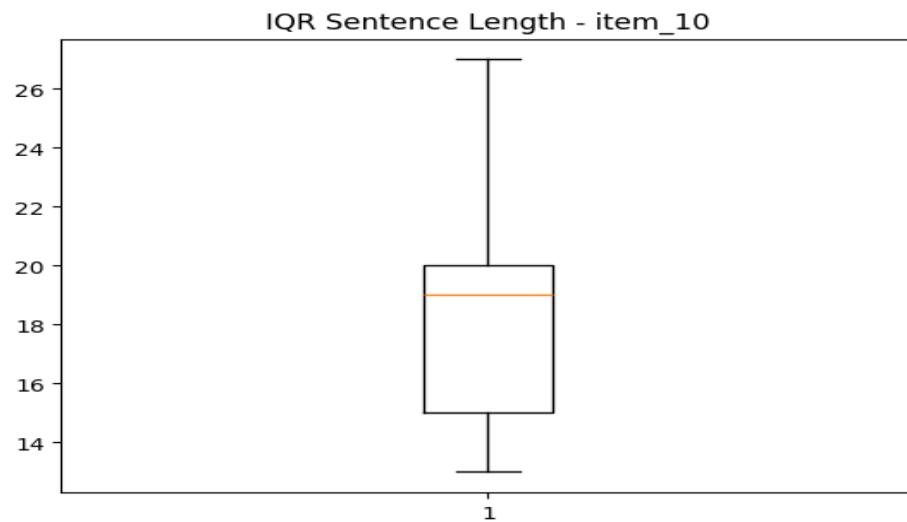
sentence_IQR_item_0.png
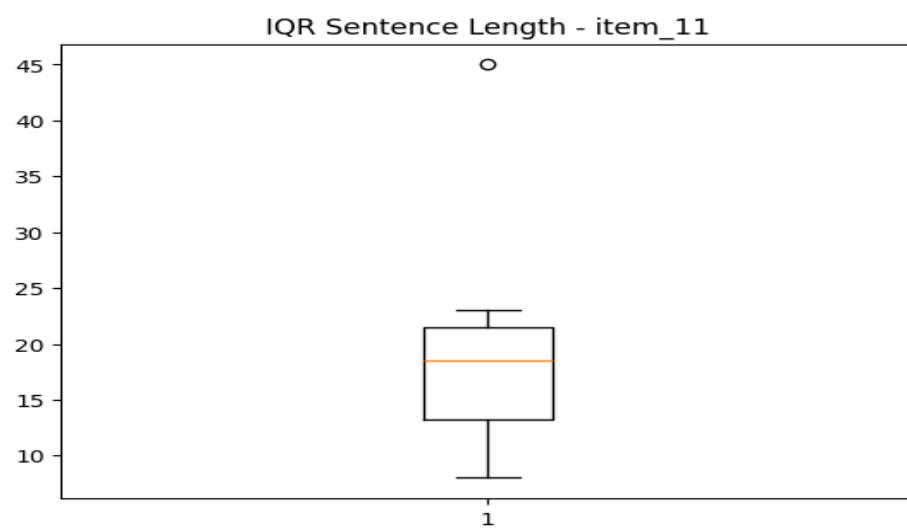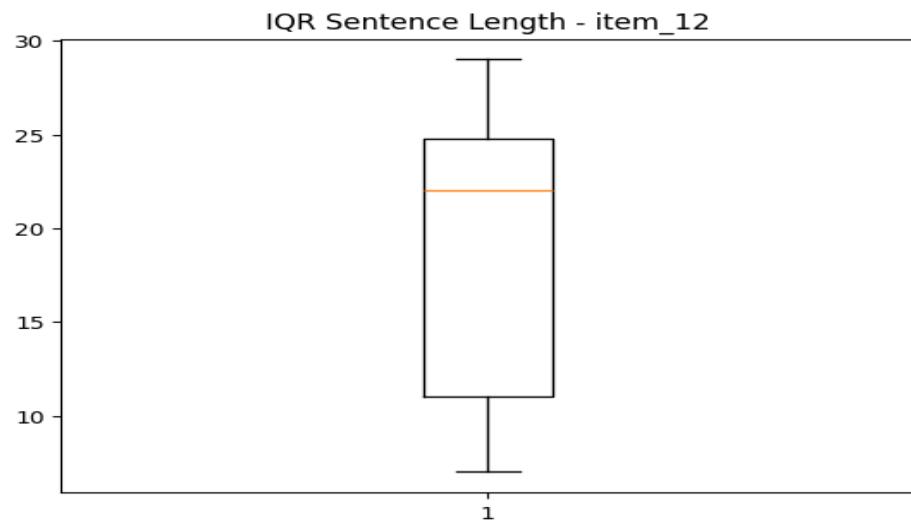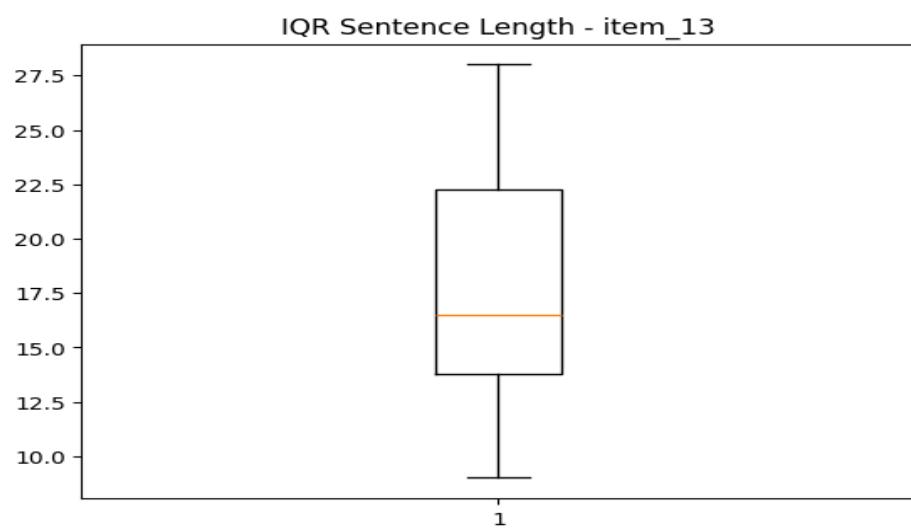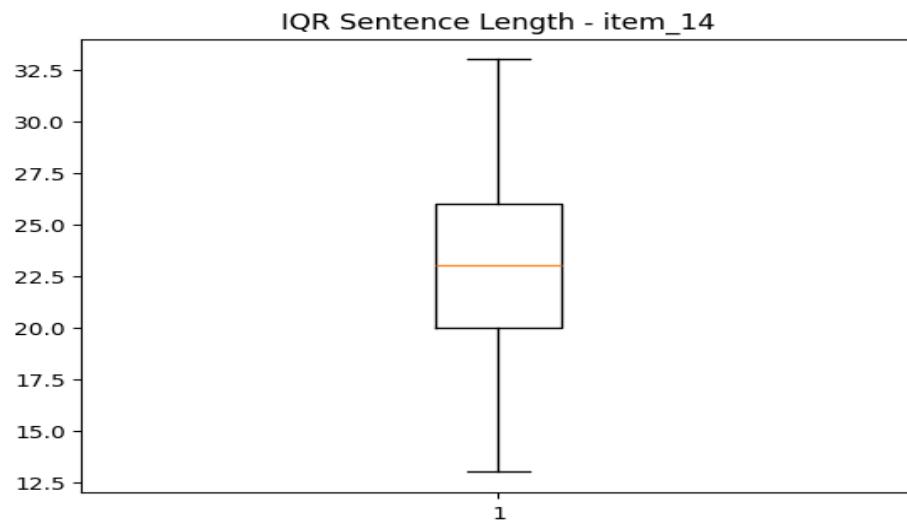
IQR Sentence Length - item_0

sentence_IQR_item_1.png

IQR Sentence Length - item_1

sentence_IQR_item_10.png

IQR Sentence Length - item_10

sentence_IQR_item_11.png

IQR Sentence Length - item_11

sentence_IQR_item_12.png

IQR Sentence Length - item_12

sentence_IQR_item_13.png

IQR Sentence Length - item_13

sentence_IQR_item_14.png

IQR Sentence Length - item_14

sentence_IQR_item_15.png

IQR Sentence Length - item_15

sentence_IQR_item_16.png

IQR Sentence Length - item_16

sentence_IQR_item_17.png

IQR Sentence Length - item_17

sentence_IQR_item_18.png

IQR Sentence Length - item_18

*sentence_IQR_item_19.png*

IQR Sentence Length - item_19
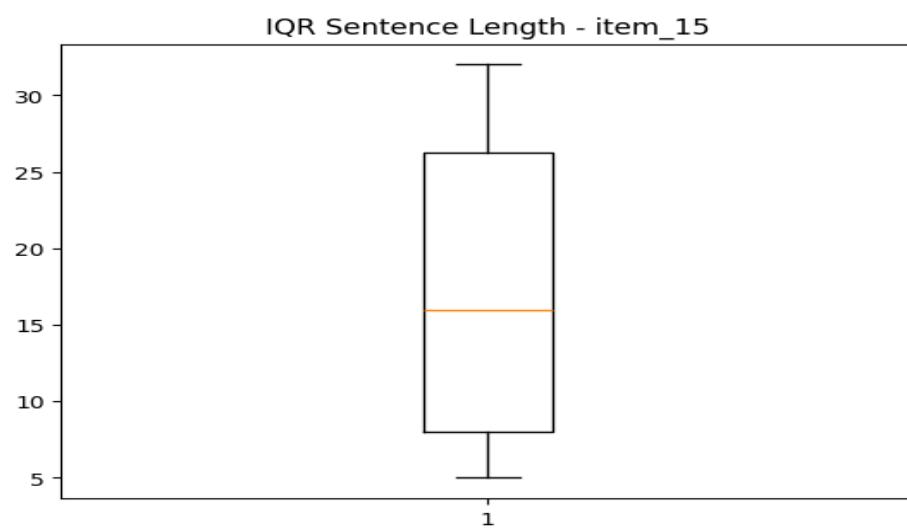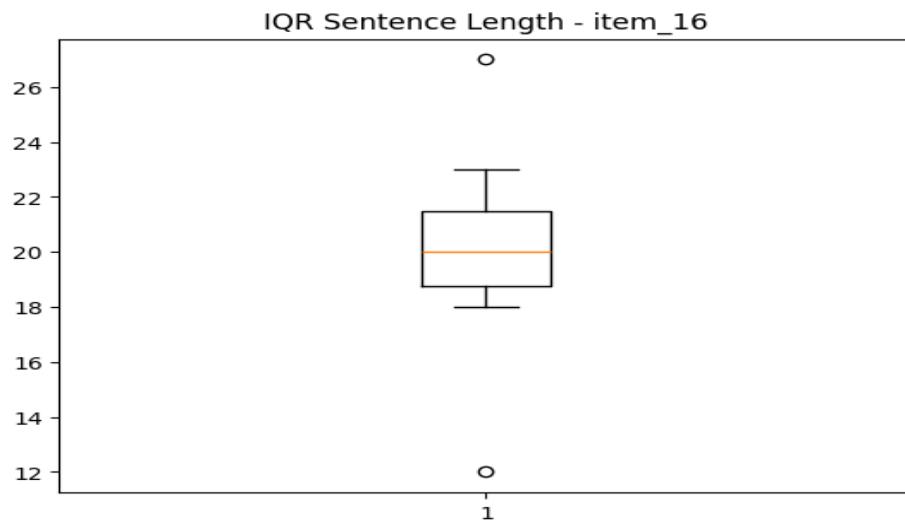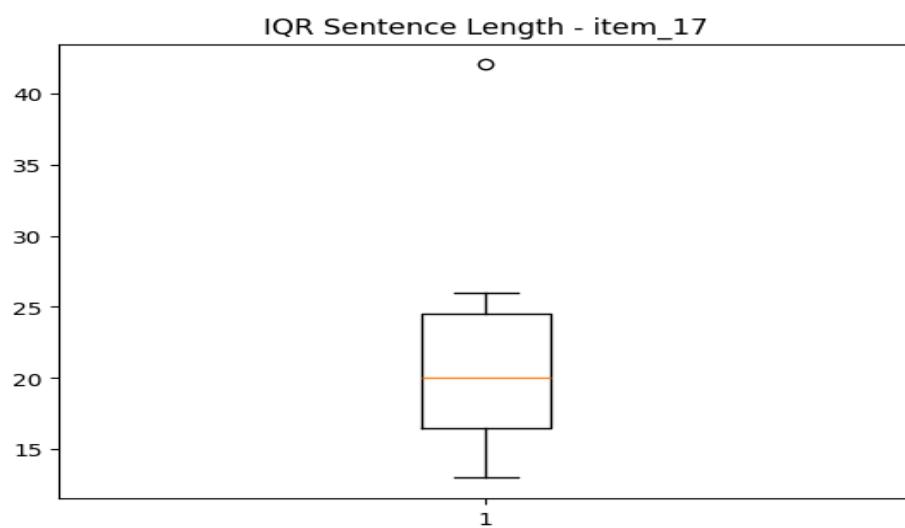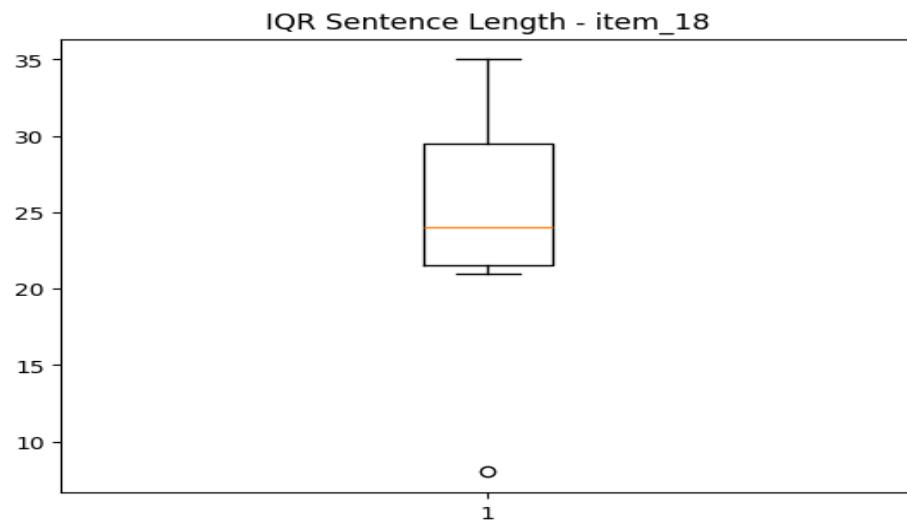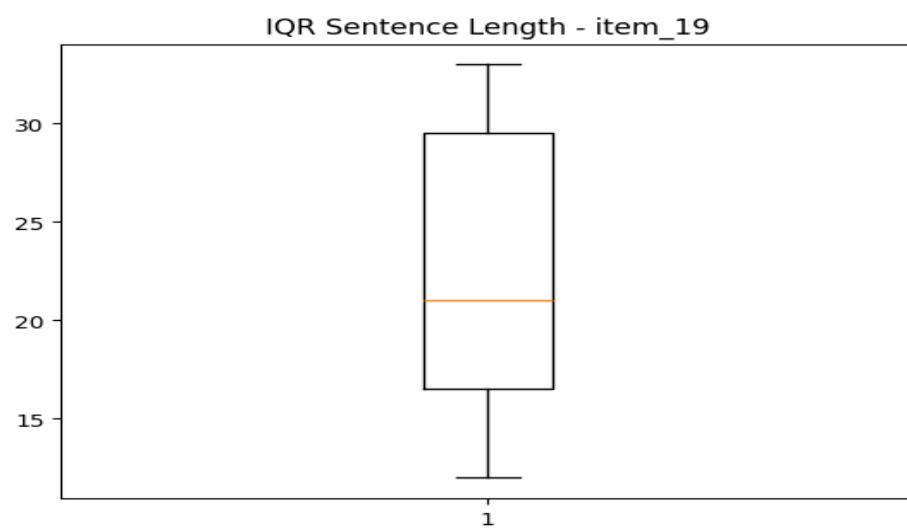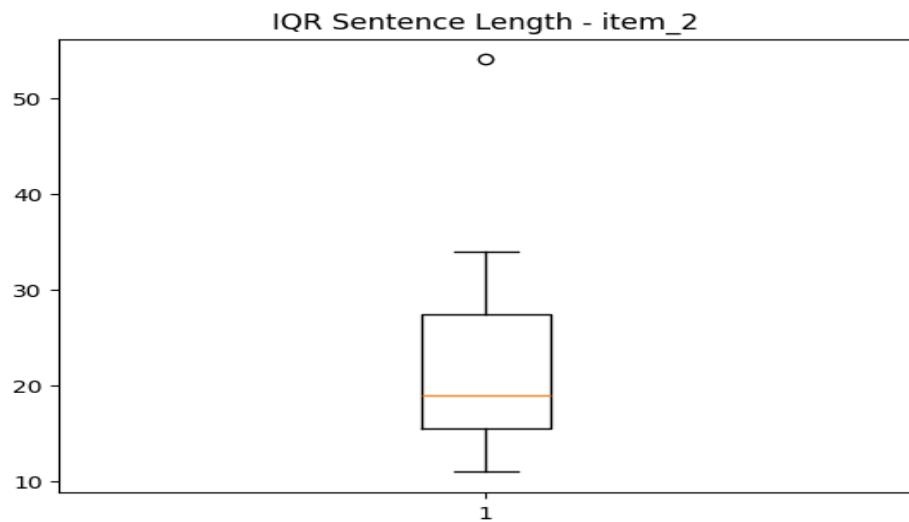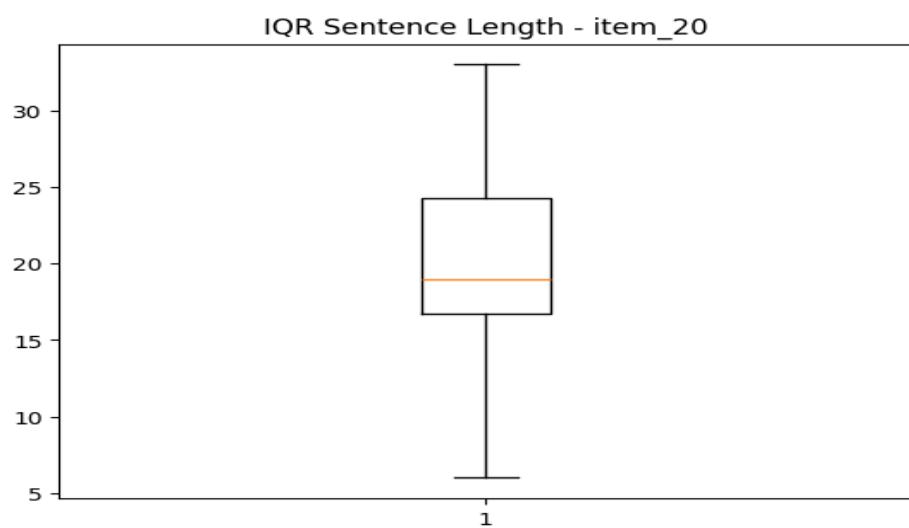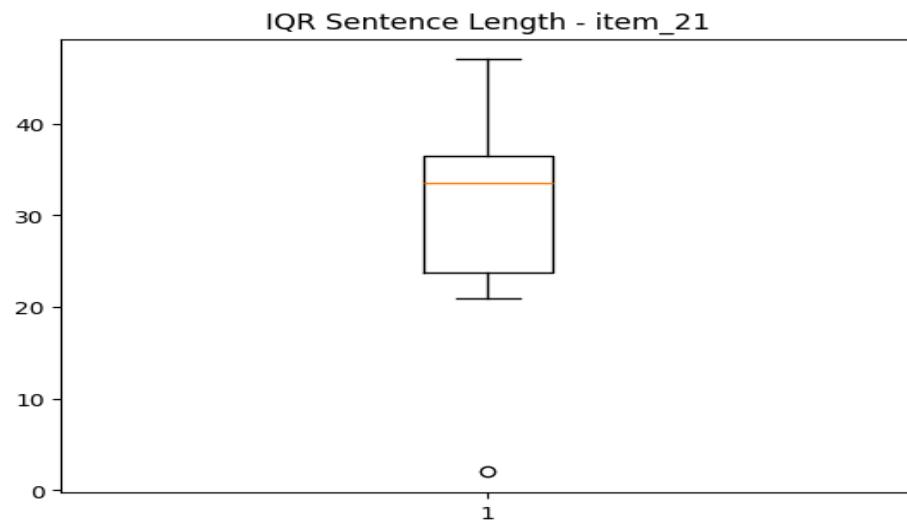
sentence_IQR_item_2.png
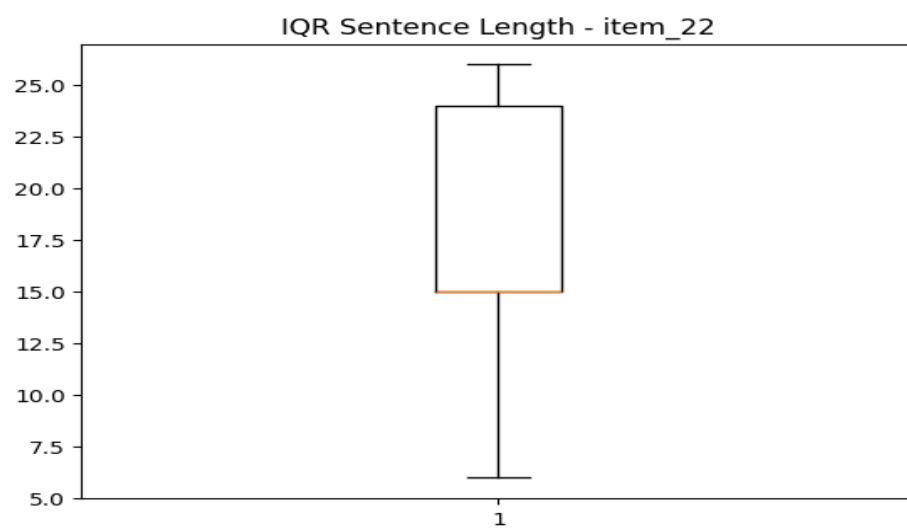
sentence_IQR_item_20.png
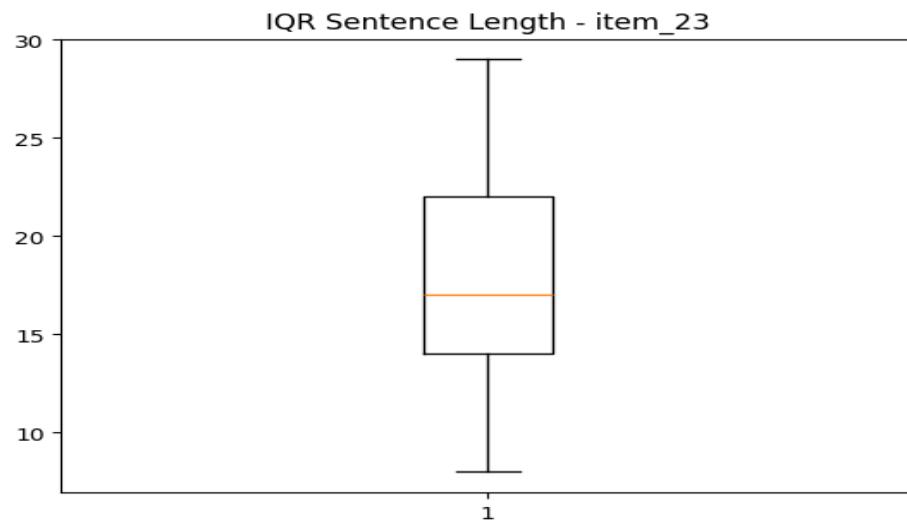
sentence_IQR_item_21.png
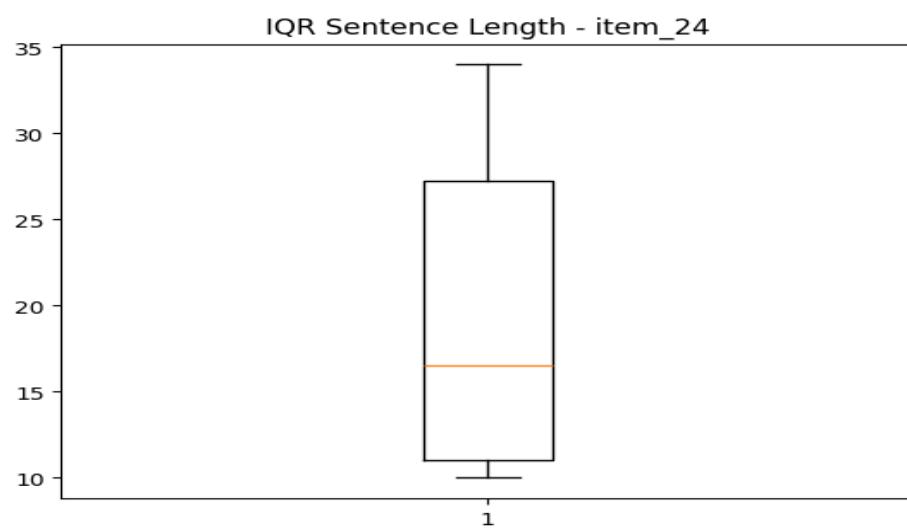
sentence_IQR_item_22.png
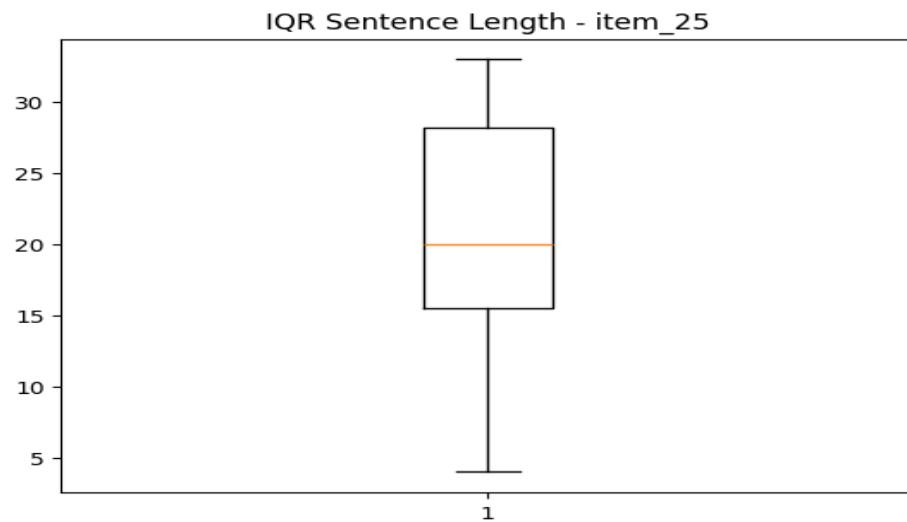
sentence_IQR_item_23.png
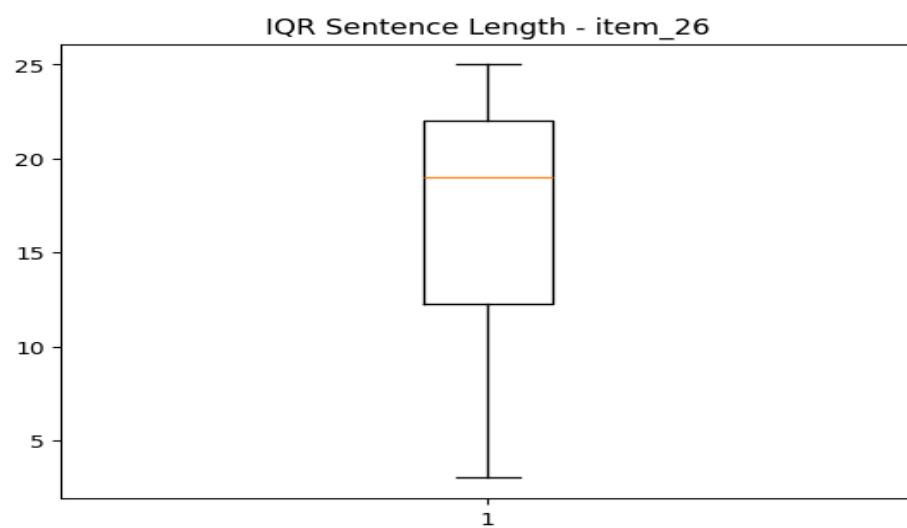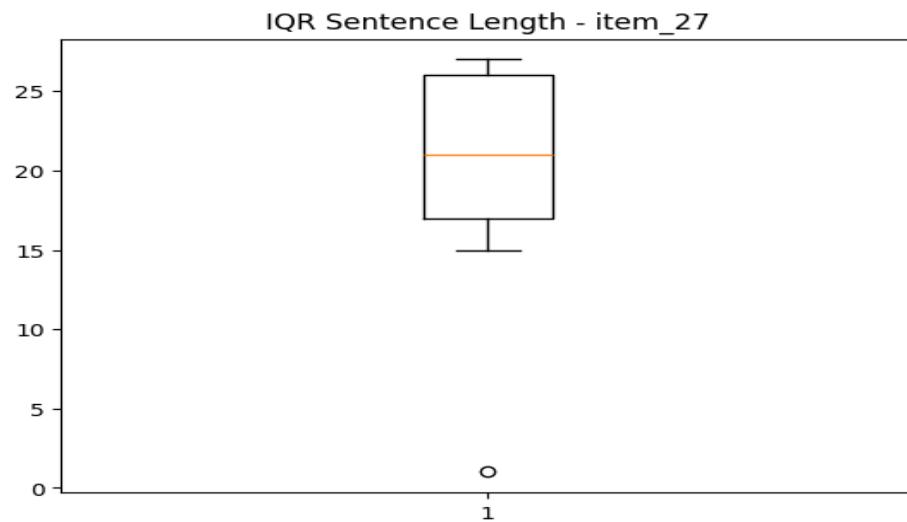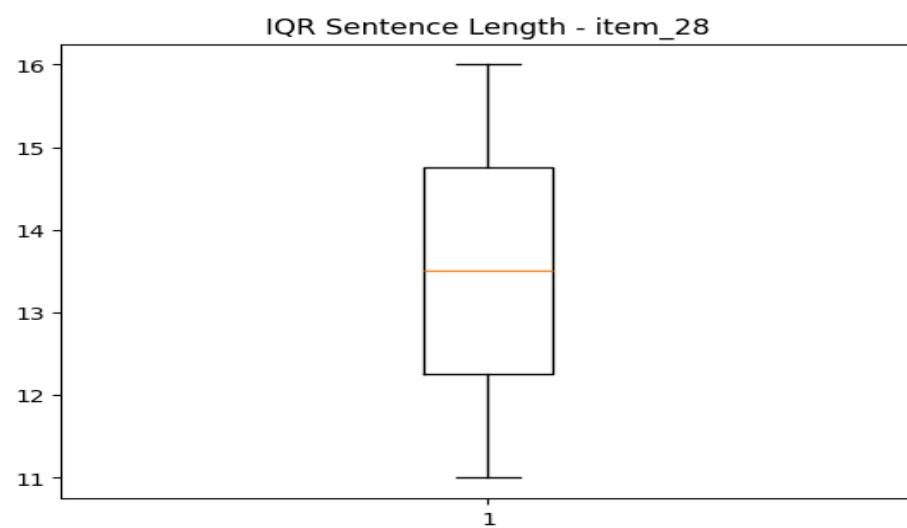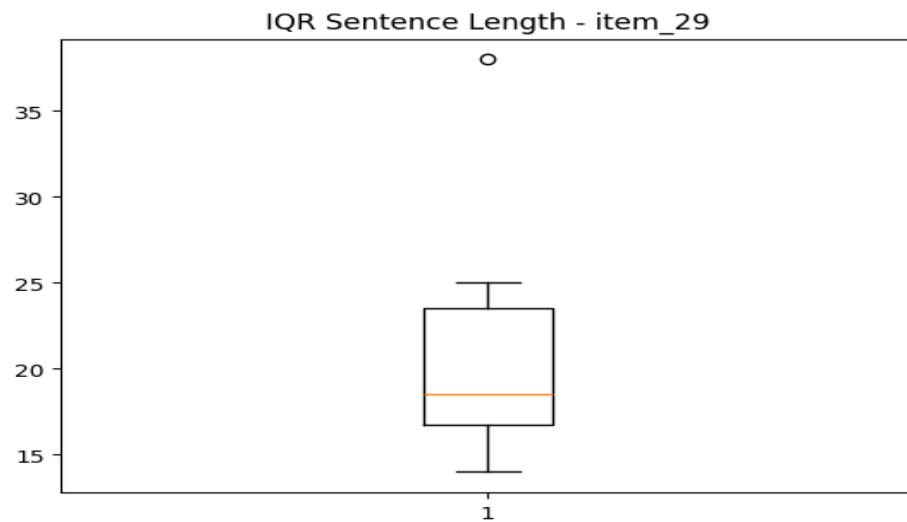
IQR Sentence Length - item_23

sentence_IQR_item_24.png

IQR Sentence Length - item_24

sentence_IQR_item_25.png

IQR Sentence Length - item_25

sentence_IQR_item_26.png

IQR Sentence Length - item_26

sentence_IQR_item_27.png

IQR Sentence Length - item_27

sentence_IQR_item_28.png

IQR Sentence Length - item_28

sentence_IQR_item_29.png

IQR Sentence Length - item_29

sentence_IQR_item_3.png

IQR Sentence Length - item_3

sentence_IQR_item_30.png

IQR Sentence Length - item_30

sentence_IQR_item_31.png

IQR Sentence Length - item_31

sentence_IQR_item_32.png

IQR Sentence Length - item_32

sentence_IQR_item_33.png

IQR Sentence Length - item_33

sentence_IQR_item_34.png

IQR Sentence Length - item_34

sentence_IQR_item_35.png

IQR Sentence Length - item_35

sentence_IQR_item_36.png

IQR Sentence Length - item_36

sentence_IQR_item_37.png

IQR Sentence Length - item_37
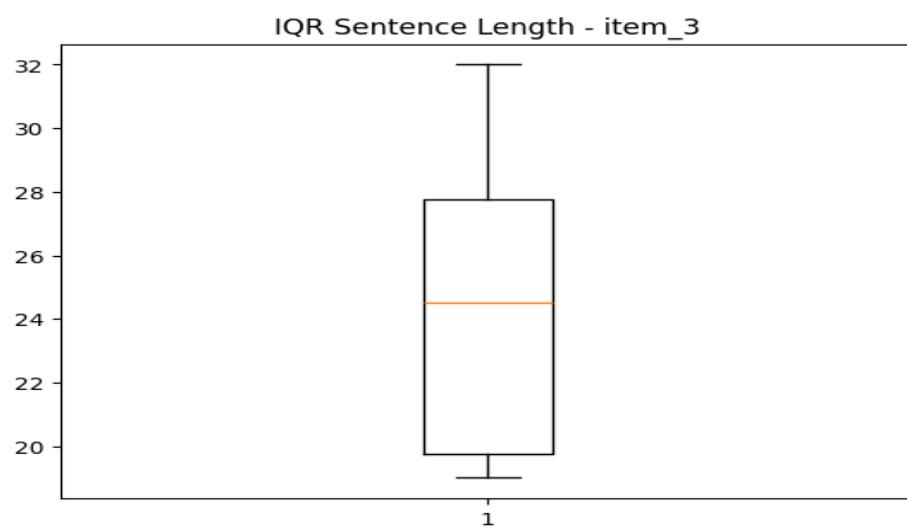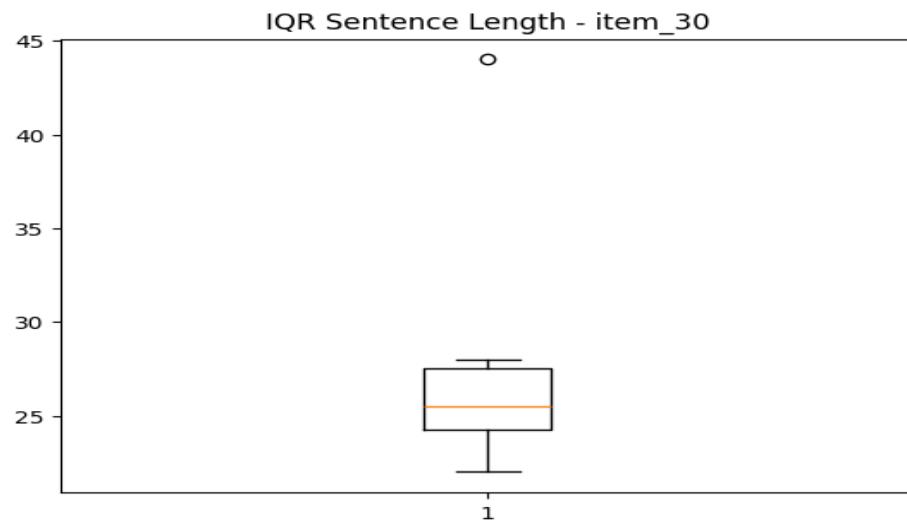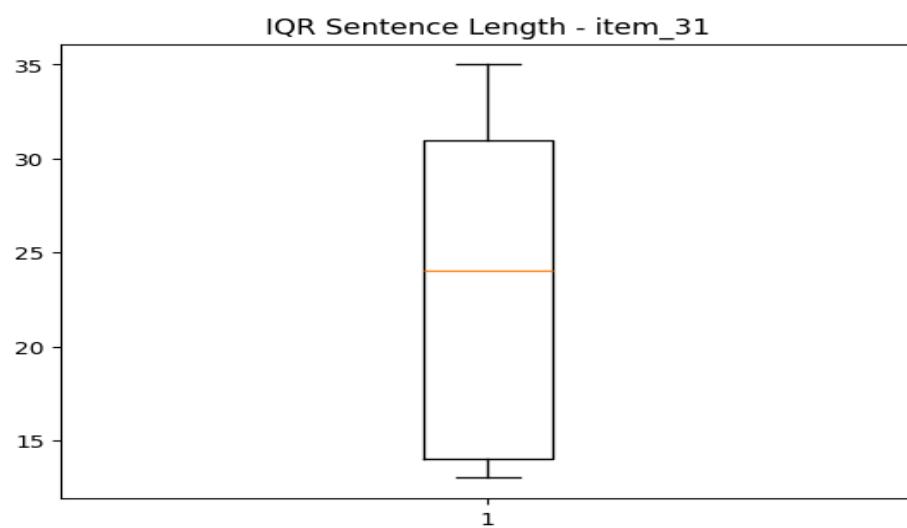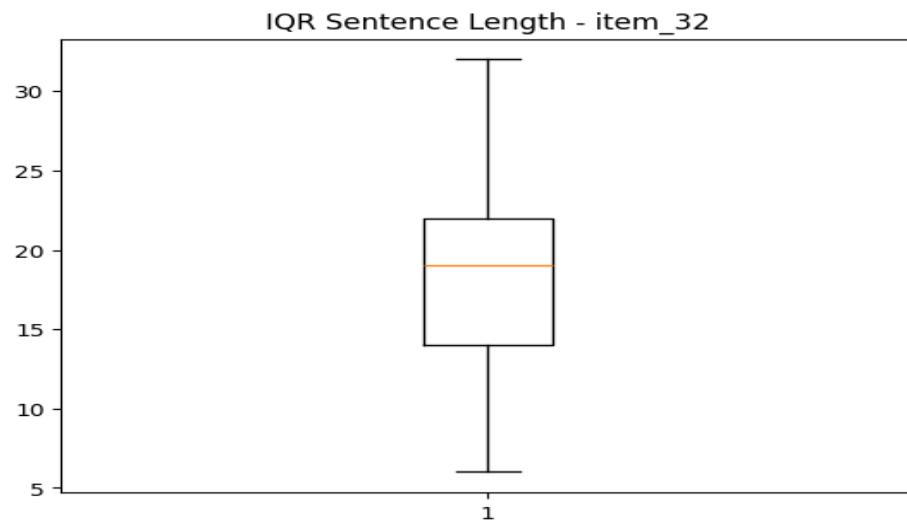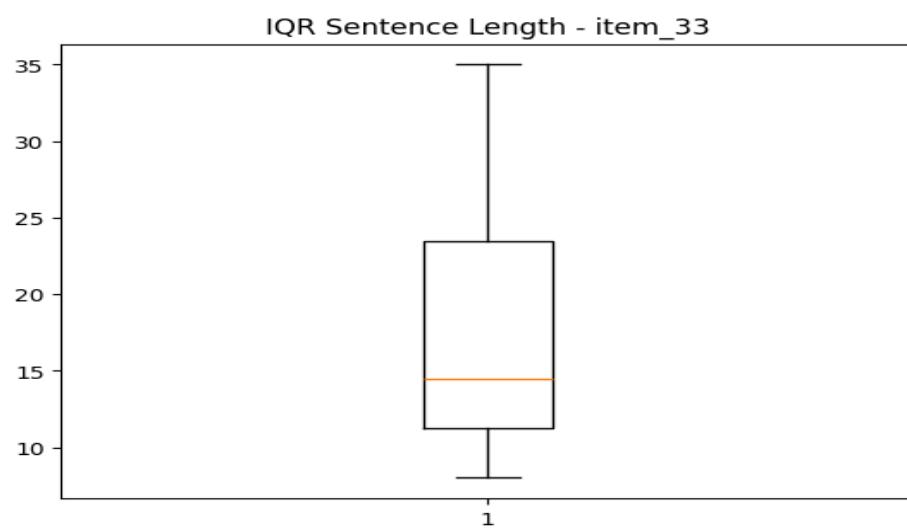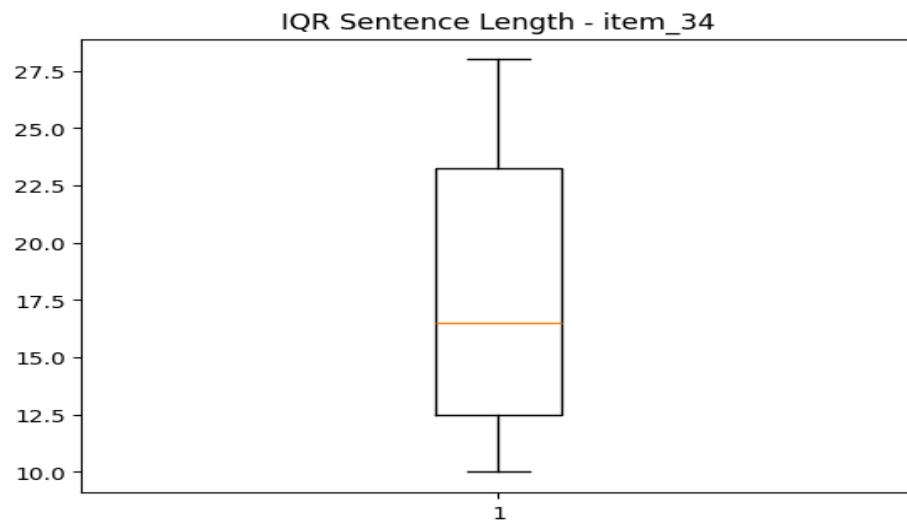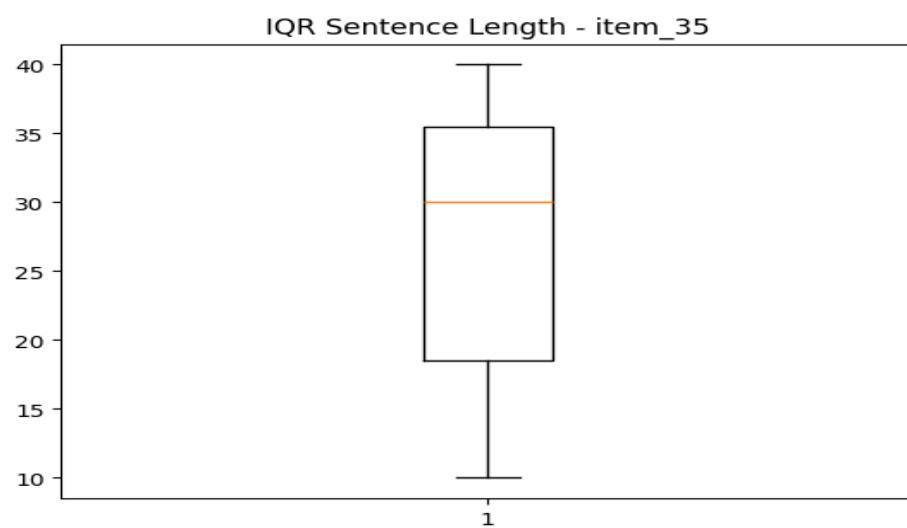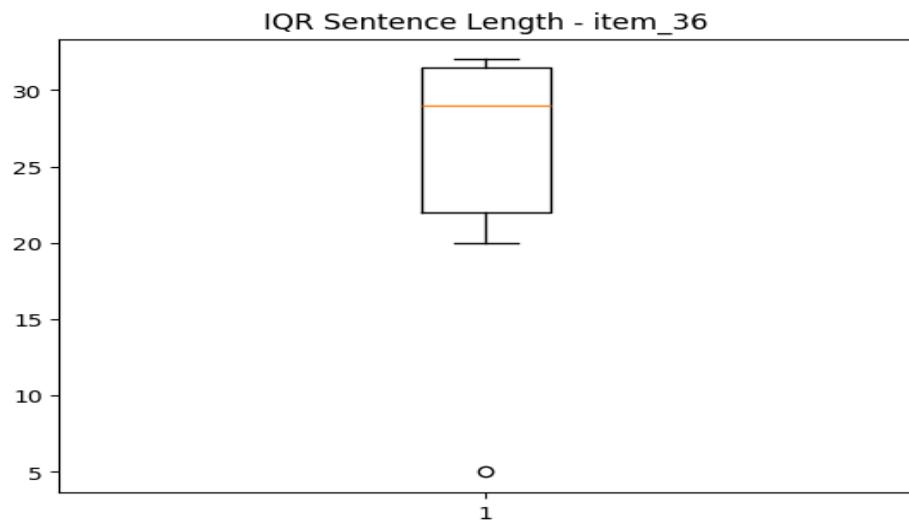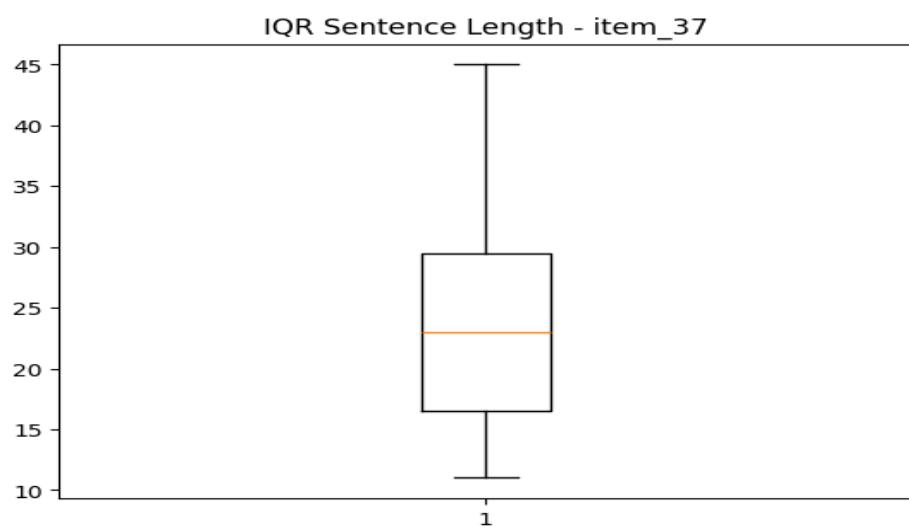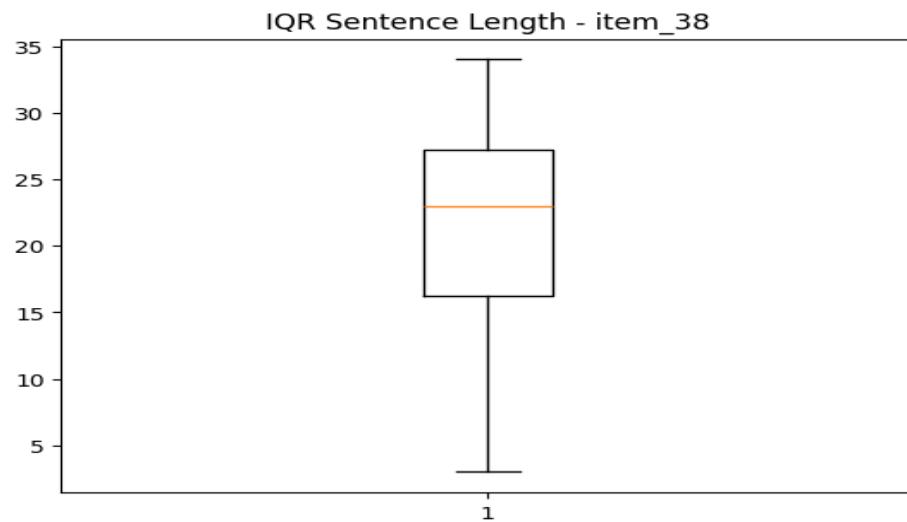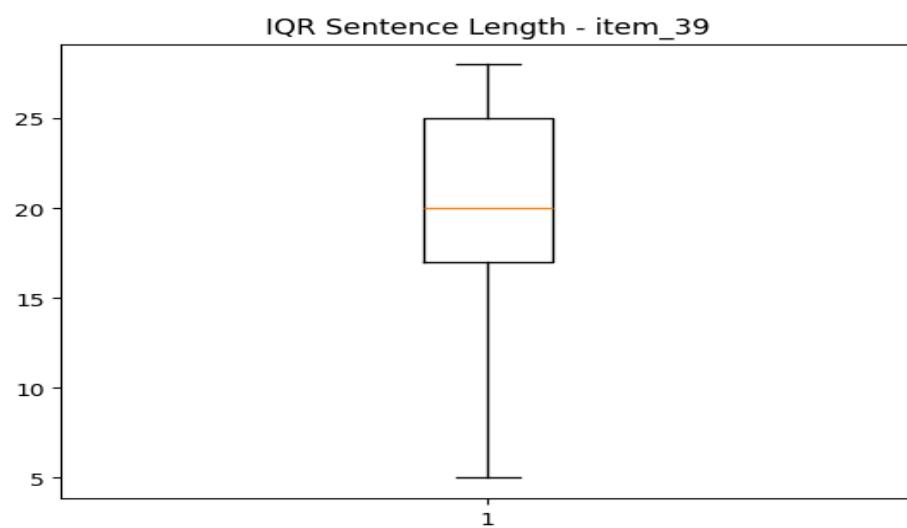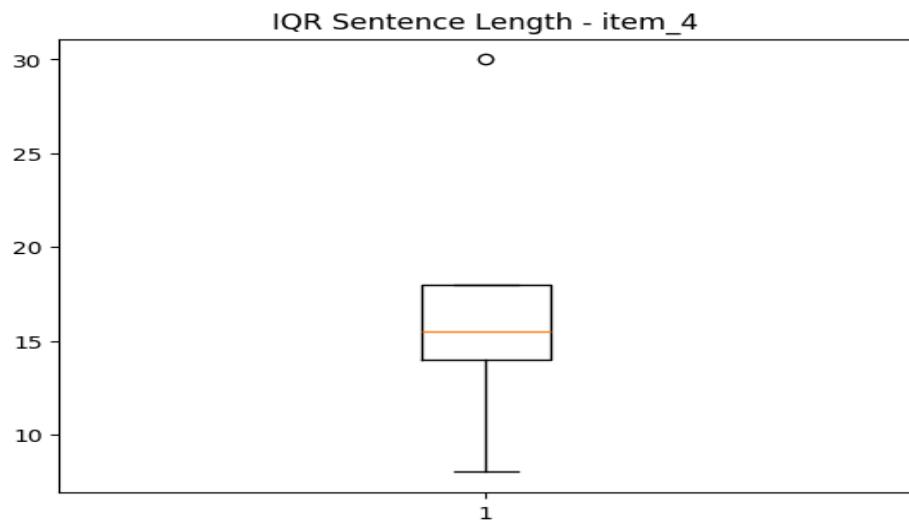
sentence_IQR_item_38.png
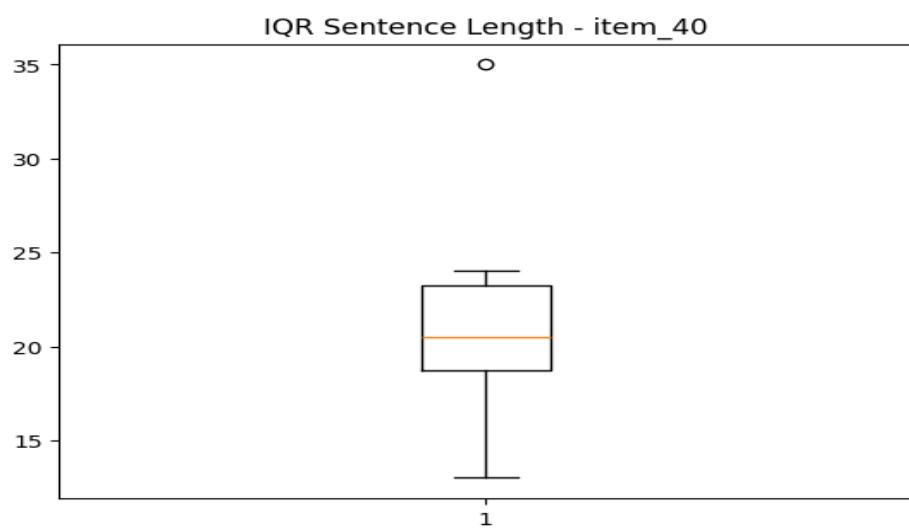
sentence_IQR_item_39.png
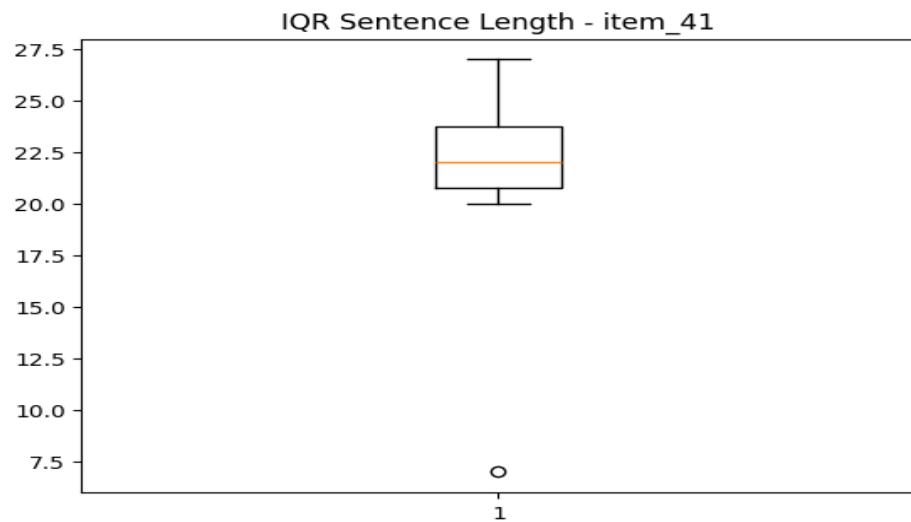
sentence_IQR_item_4.png
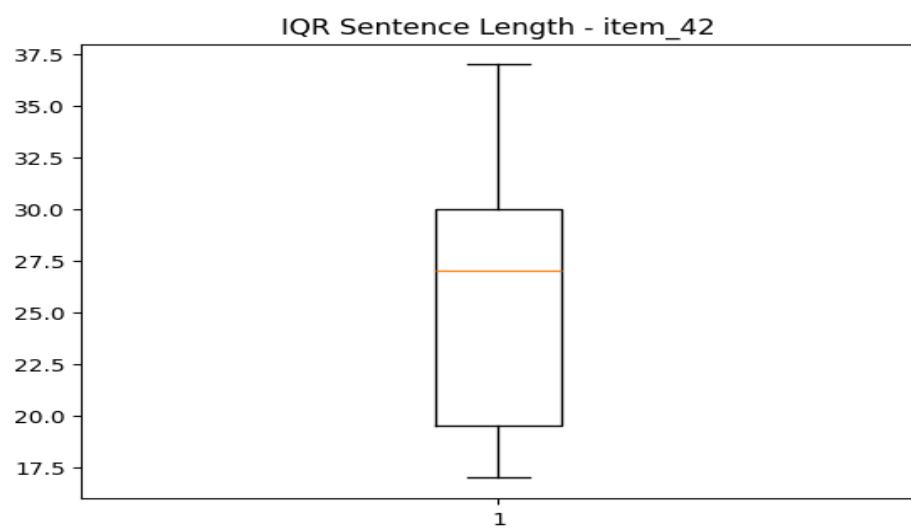
IQR Sentence Length - item_4

sentence_IQR_item_40.png

IQR Sentence Length - item_40

sentence_IQR_item_41.png

IQR Sentence Length - item_41

sentence_IQR_item_42.png

IQR Sentence Length - item_42
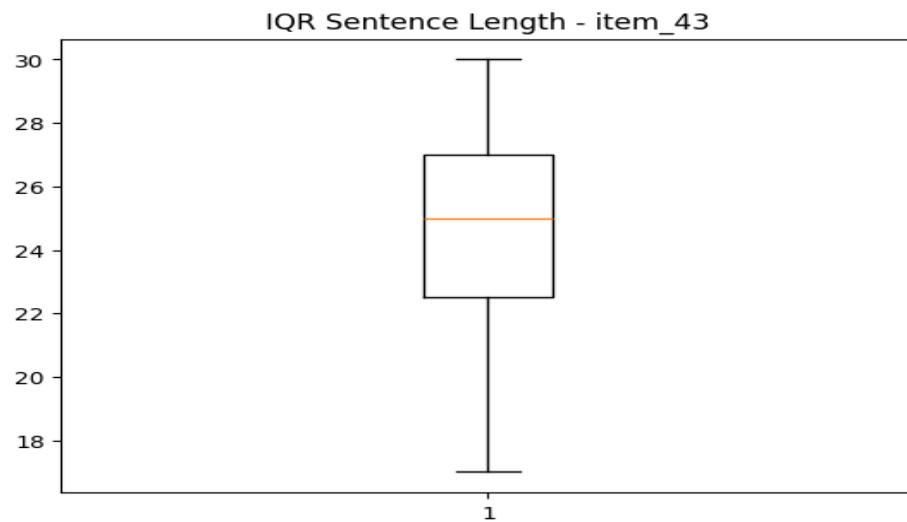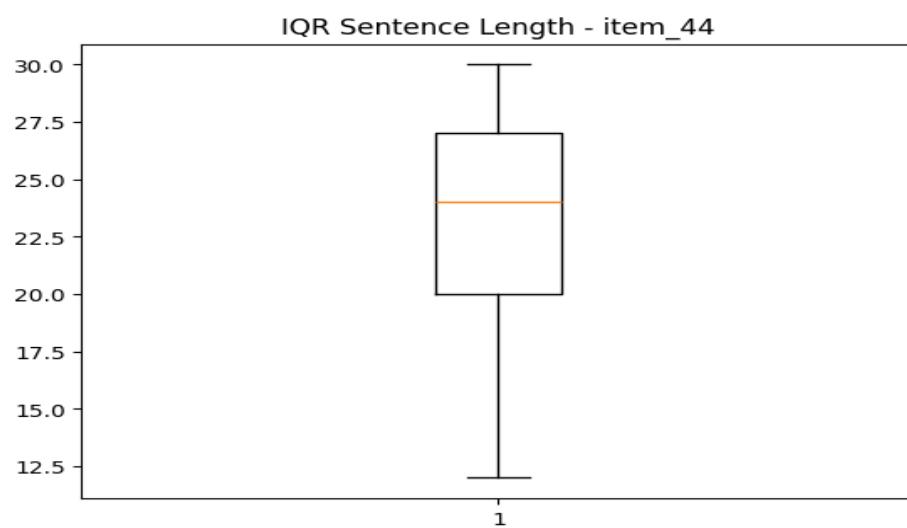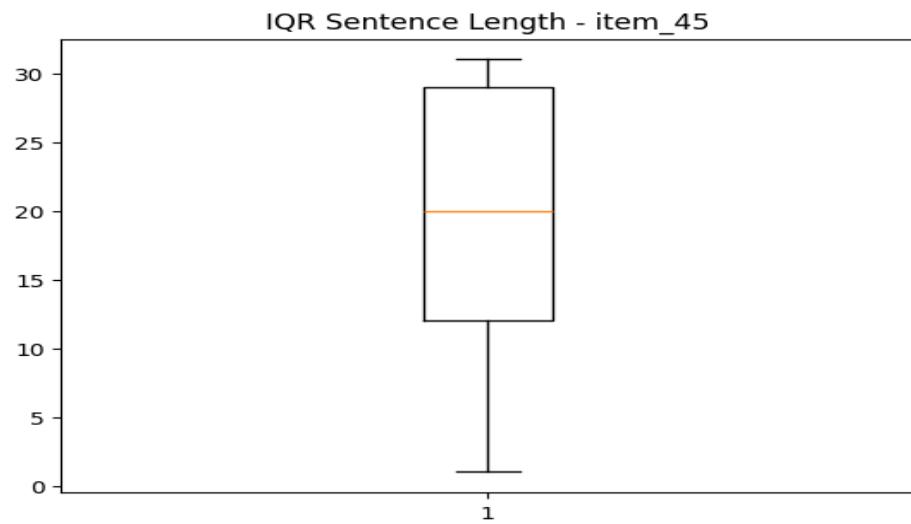
sentence_IQR_item_43.png
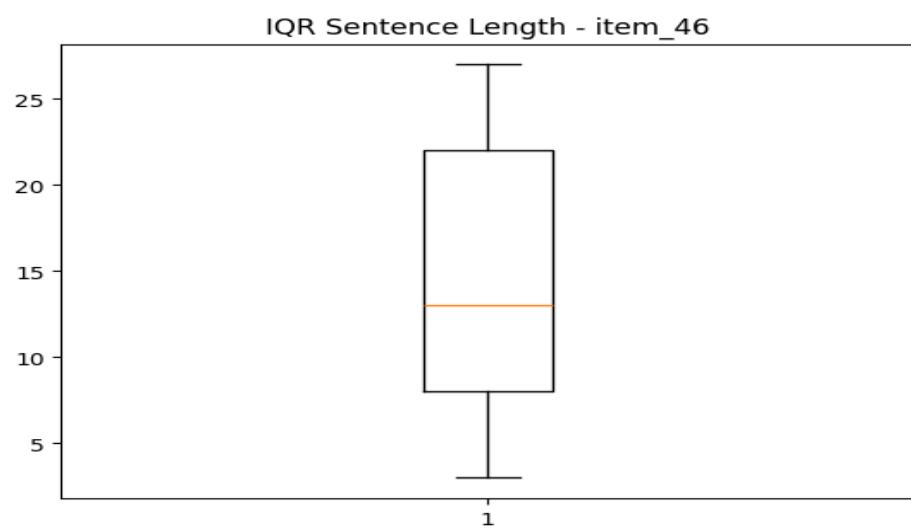
sentence_IQR_item_44.png
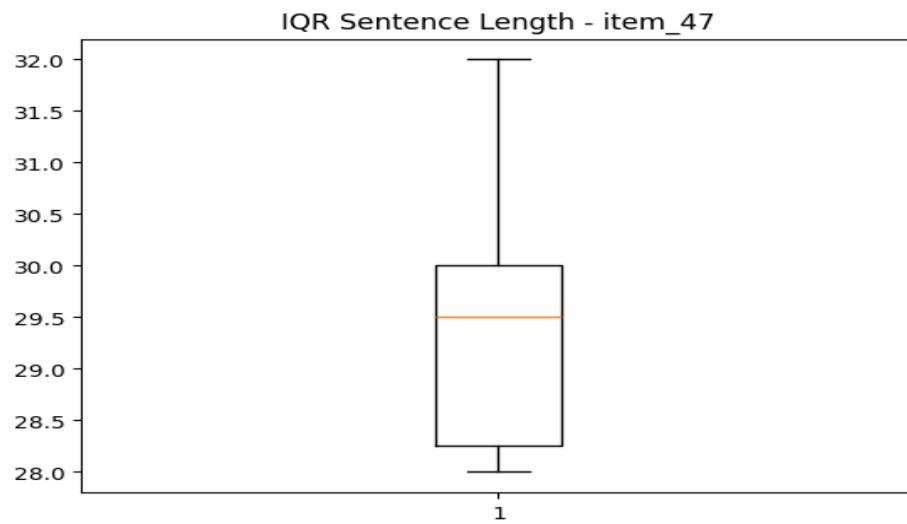
sentence_IQR_item_45.png
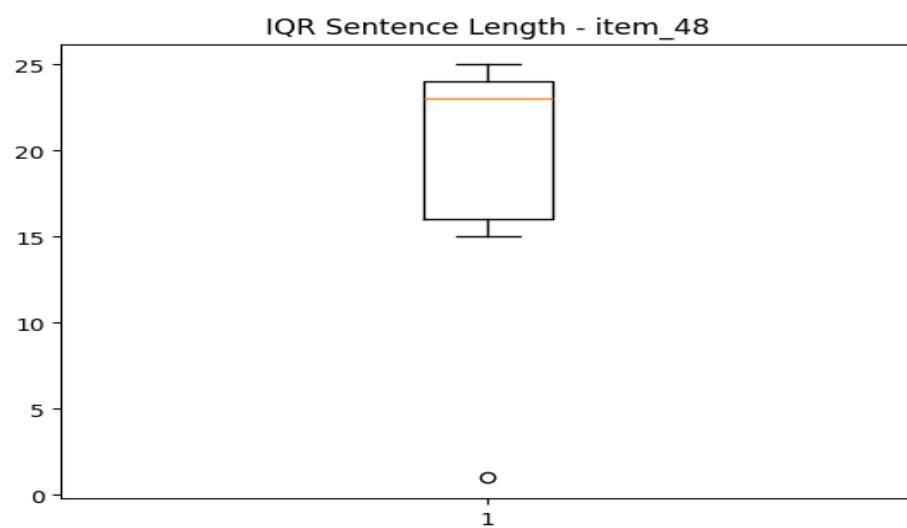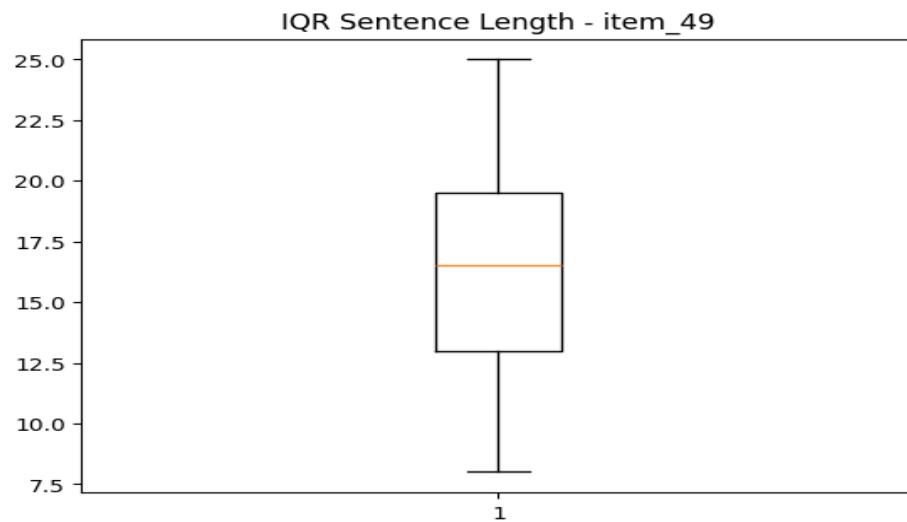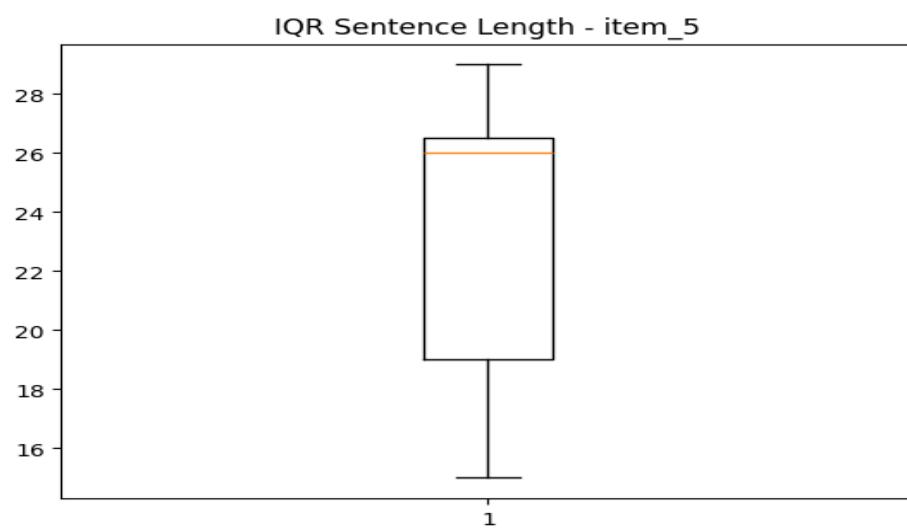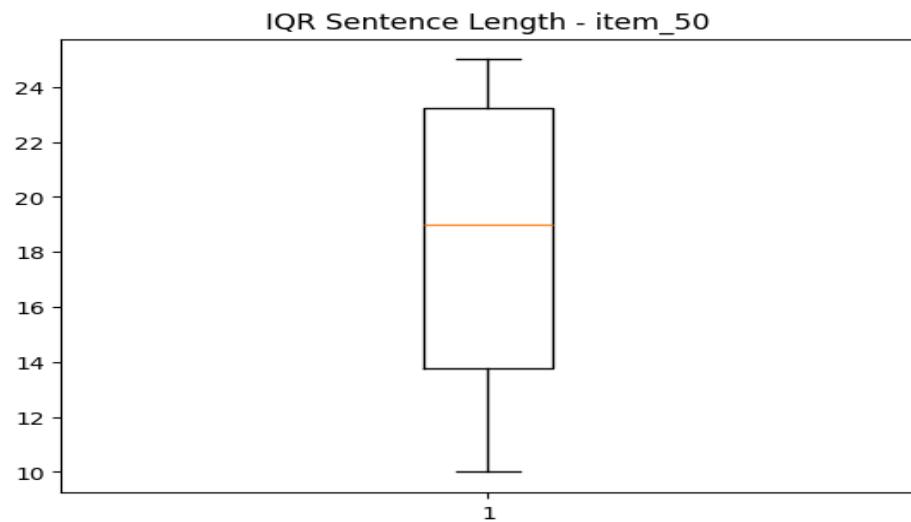
IQR Sentence Length - item_45

sentence_IQR_item_46.png

IQR Sentence Length - item_46

sentence_IQR_item_47.png

IQR Sentence Length - item_47

sentence_IQR_item_48.png

IQR Sentence Length - item_48

sentence_IQR_item_49.png

IQR Sentence Length - item_49

sentence_IQR_item_5.png

IQR Sentence Length - item_5

sentence_IQR_item_50.png

IQR Sentence Length - item_50

*sentence_IQR_item_51.png*

IQR Sentence Length - item_51
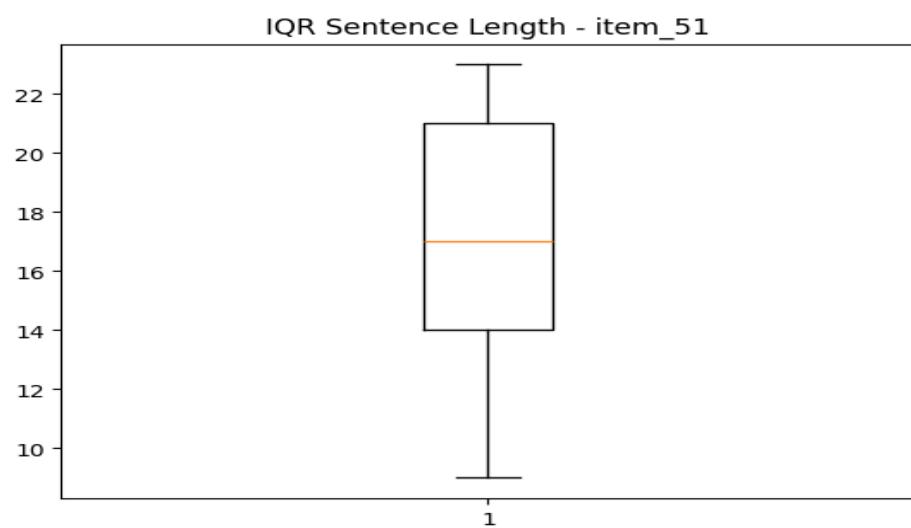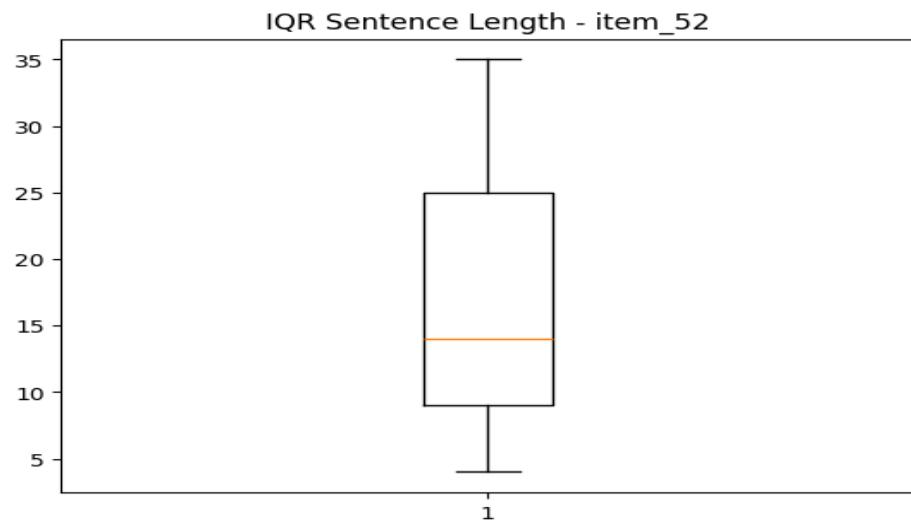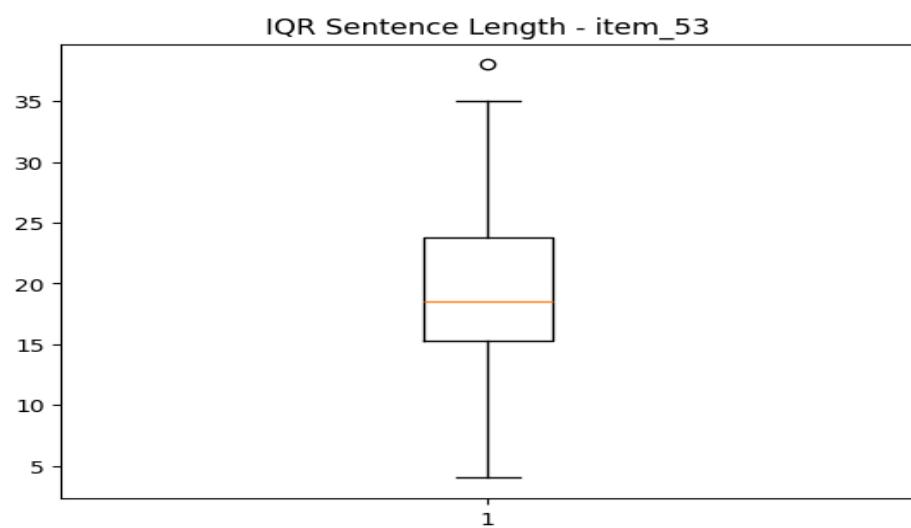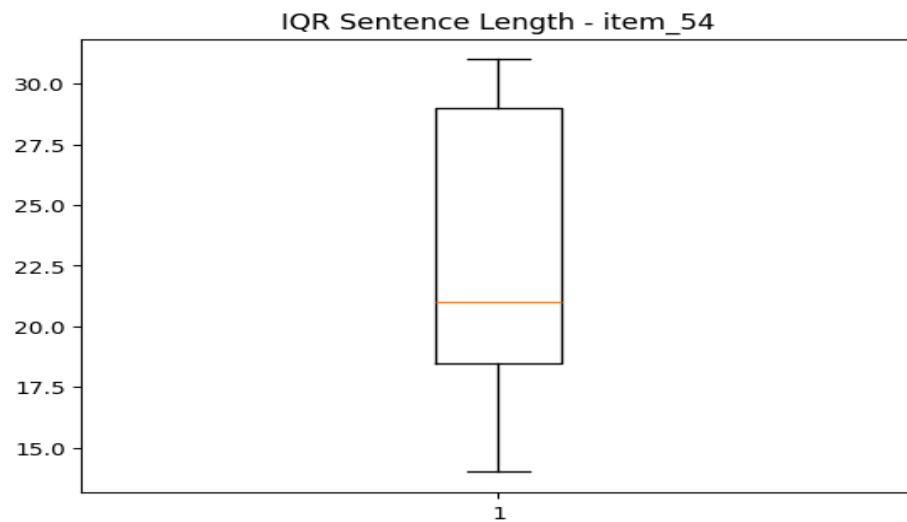
sentence_IQR_item_52.png
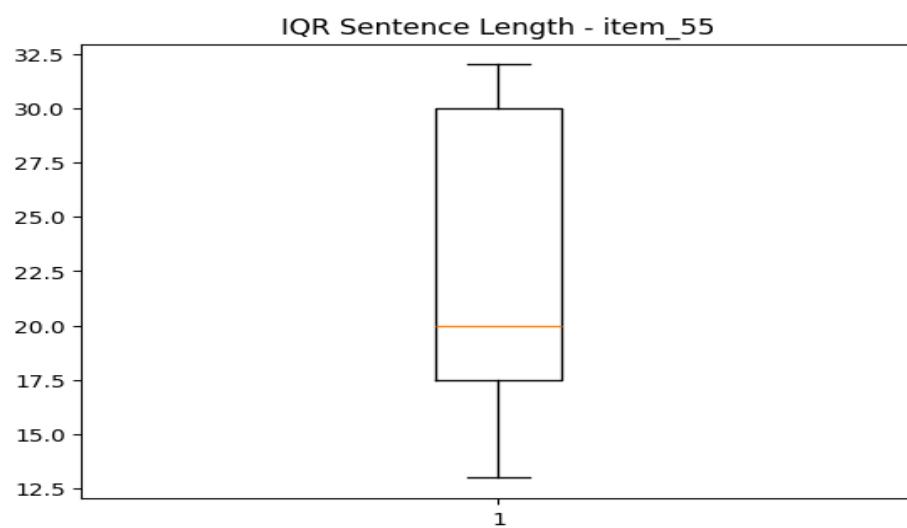
sentence_IQR_item_53.png
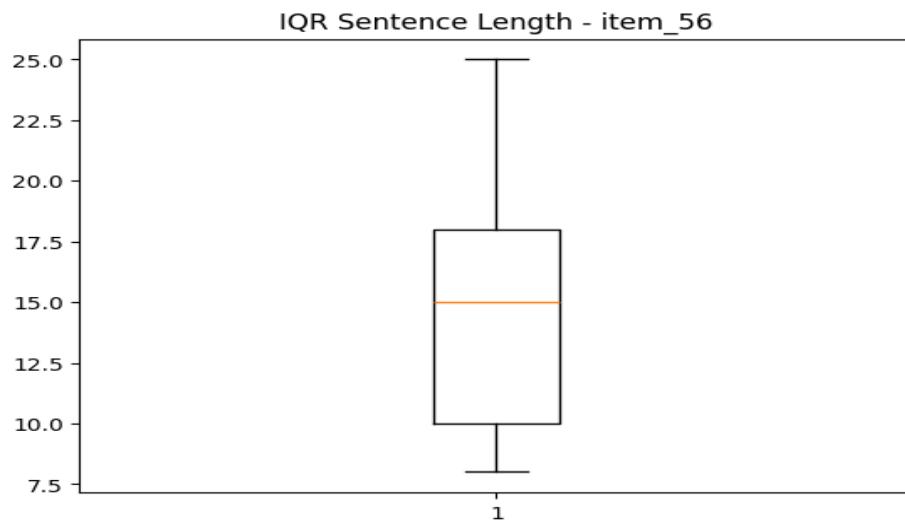
sentence_IQR_item_54.png
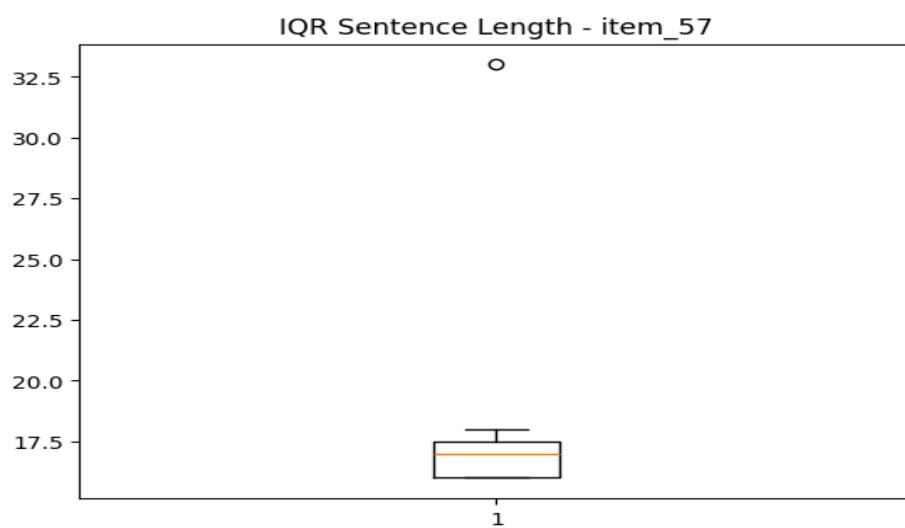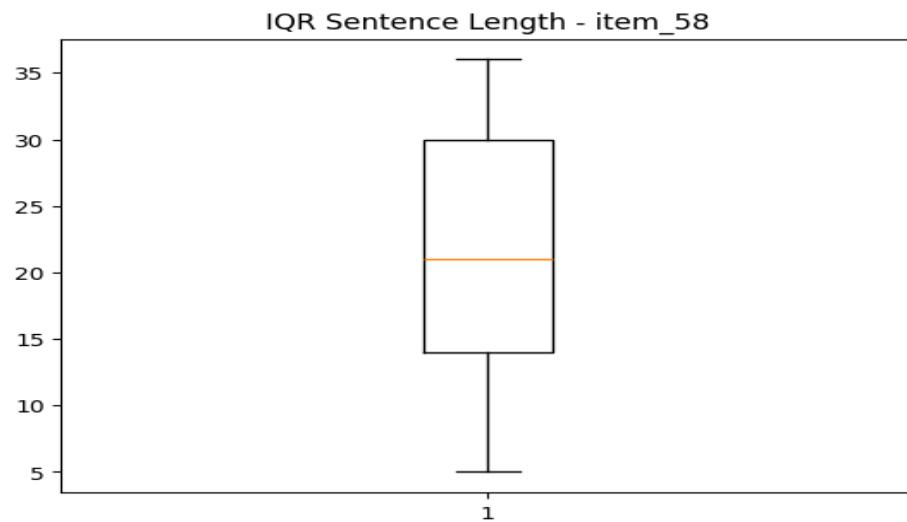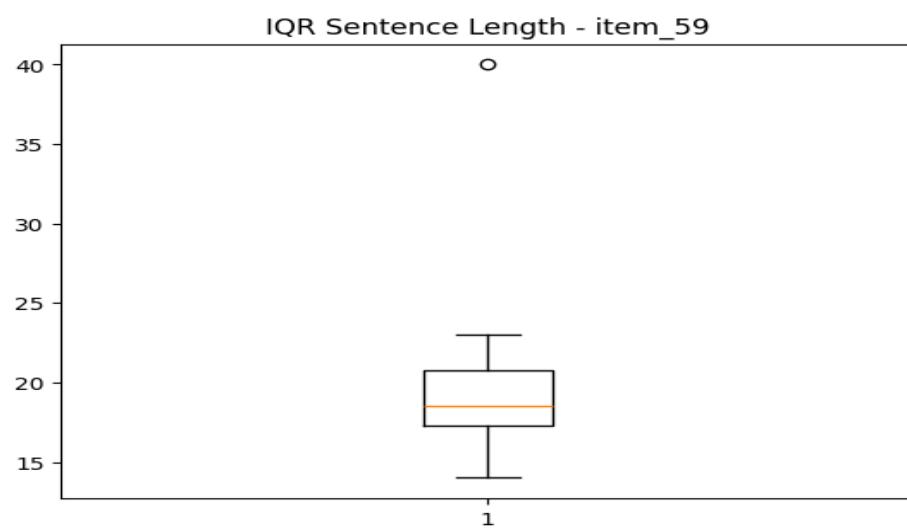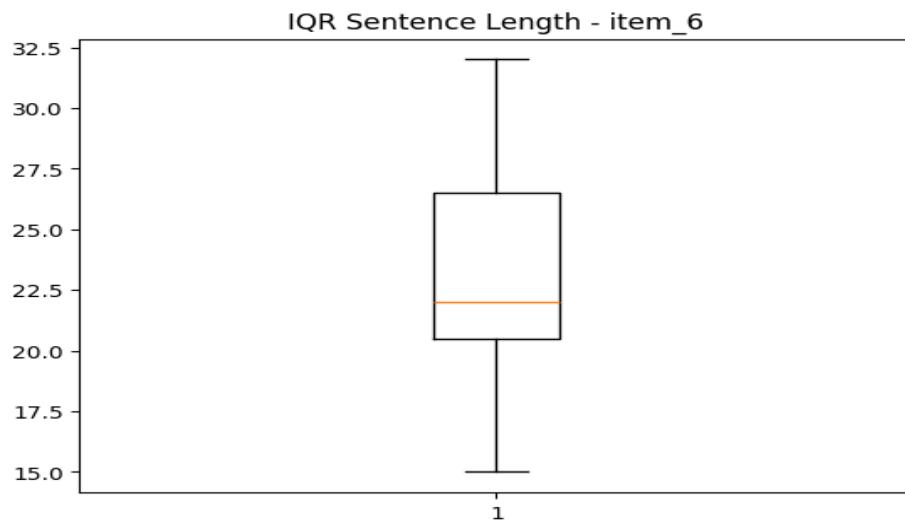

IQR Sentence Length - item_54

sentence_IQR_item_55.png

IQR Sentence Length - item_55

sentence_IQR_item_56.png

IQR Sentence Length - item_56

sentence_IQR_item_57.png

IQR Sentence Length - item_57

sentence_IQR_item_58.png

IQR Sentence Length - item_58

sentence_IQR_item_59.png

IQR Sentence Length - item_59

sentence_IQR_item_6.png

IQR Sentence Length - item_6

sentence_IQR_item_60.png

IQR Sentence Length - item_60

sentence_IQR_item_61.png

IQR Sentence Length - item_61

sentence_IQR_item_62.png

IQR Sentence Length - item_62

sentence_IQR_item_63.png

IQR Sentence Length - item_63

sentence_IQR_item_64.png

IQR Sentence Length - item_64
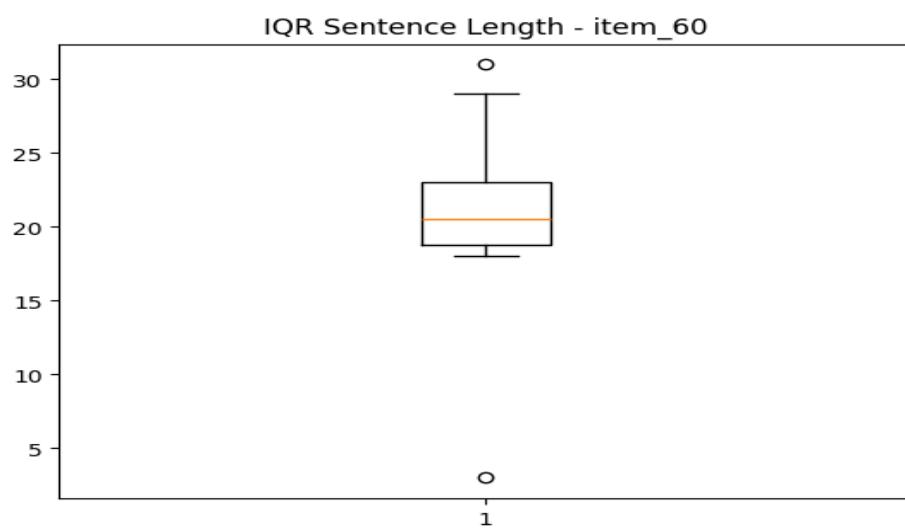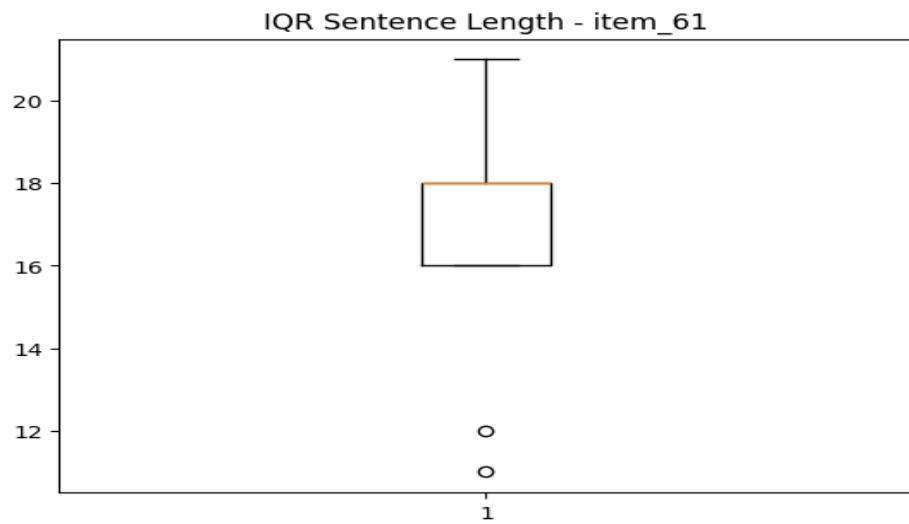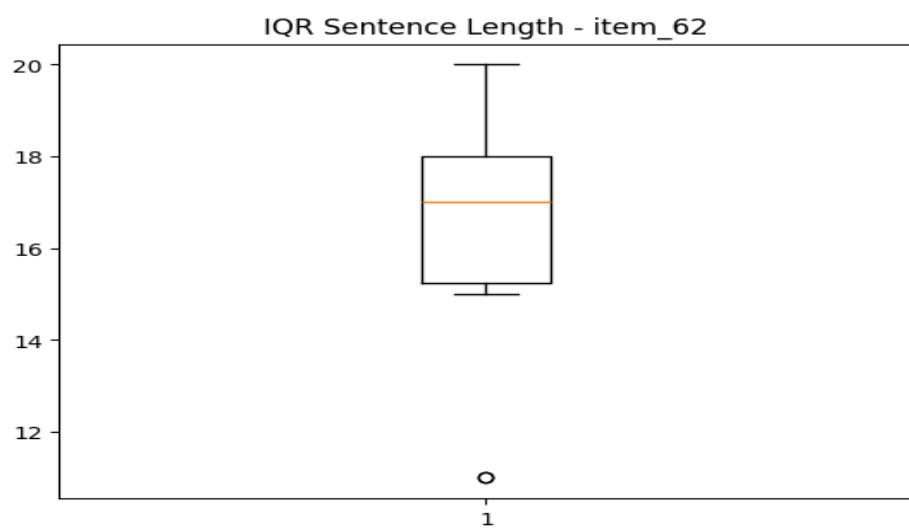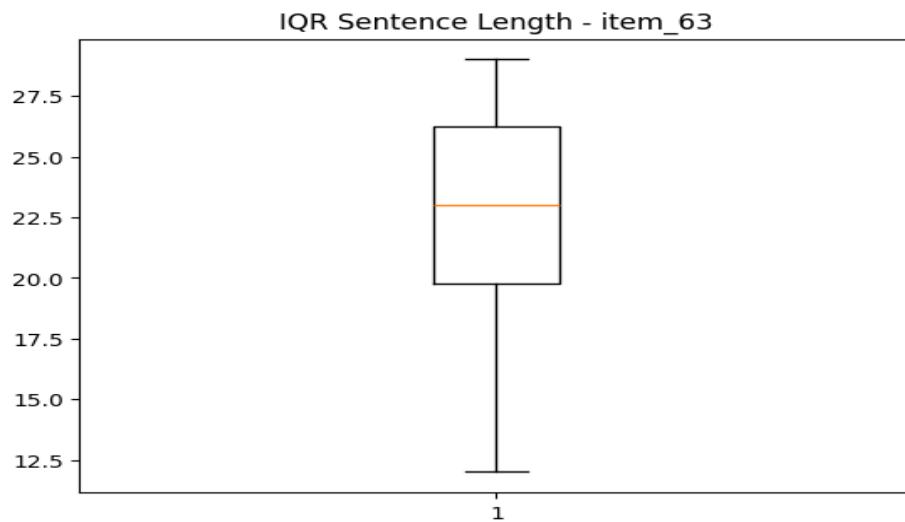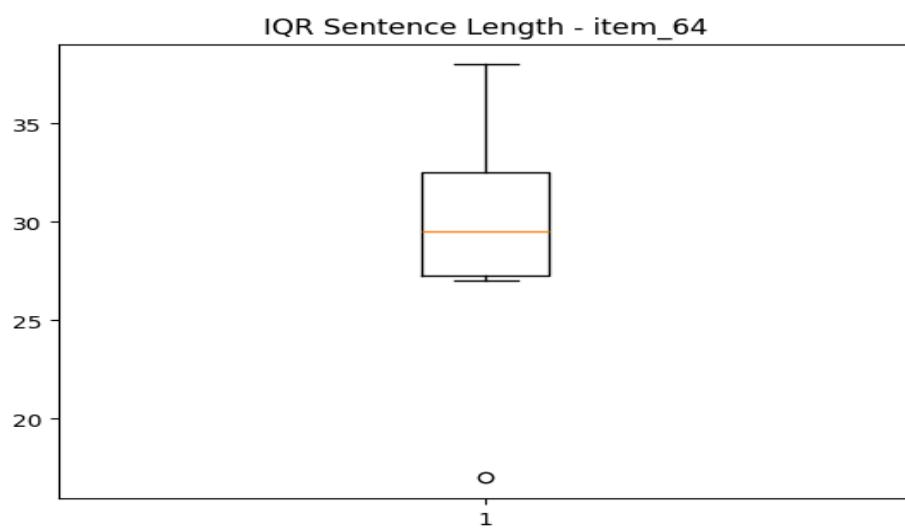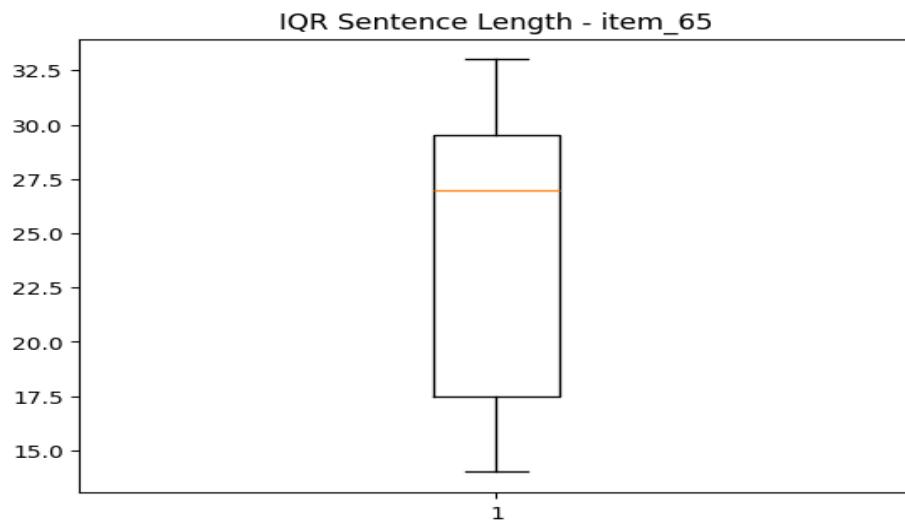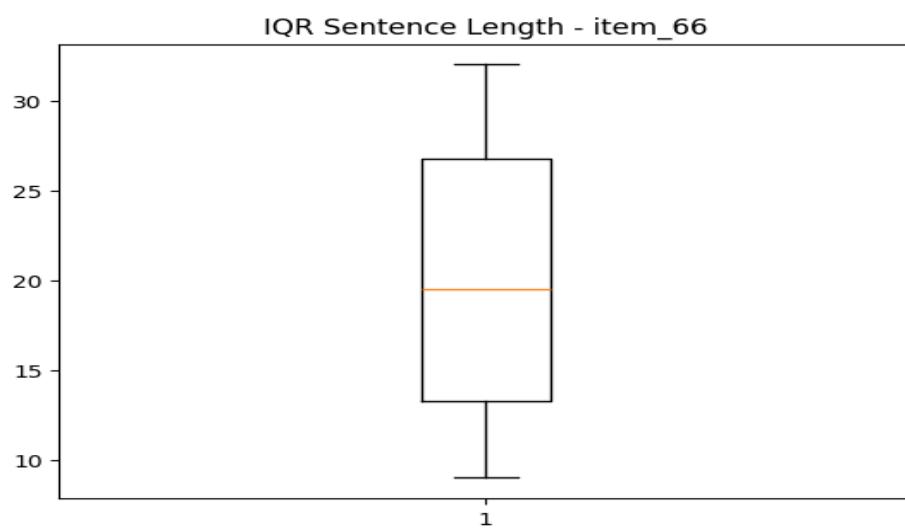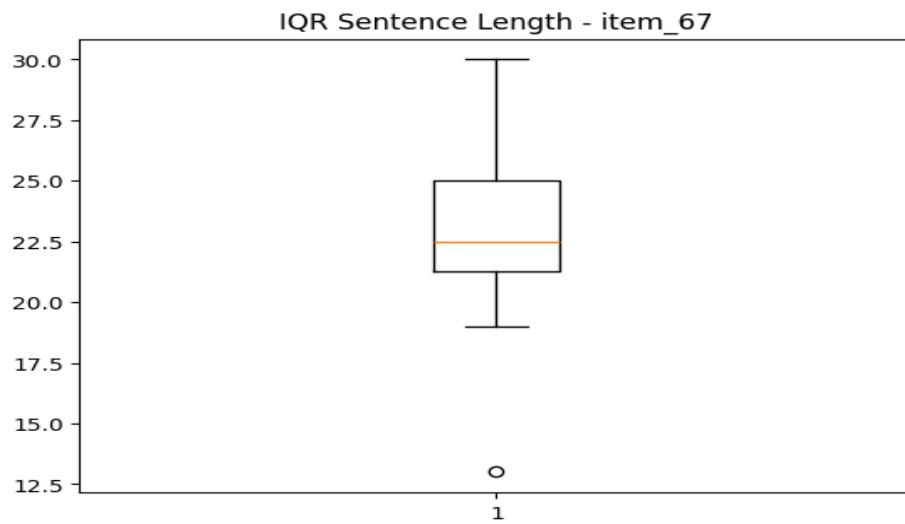
sentence_IQR_item_65.png
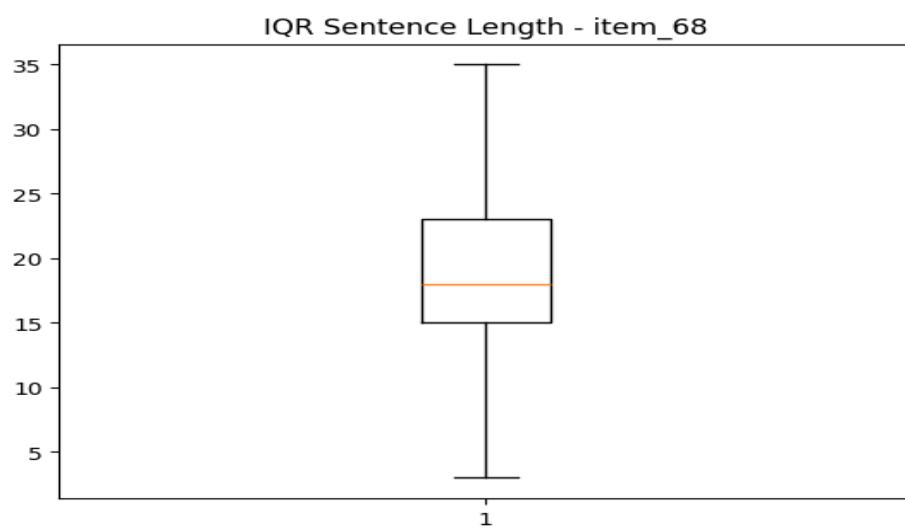

sentence_IQR_item_66.png
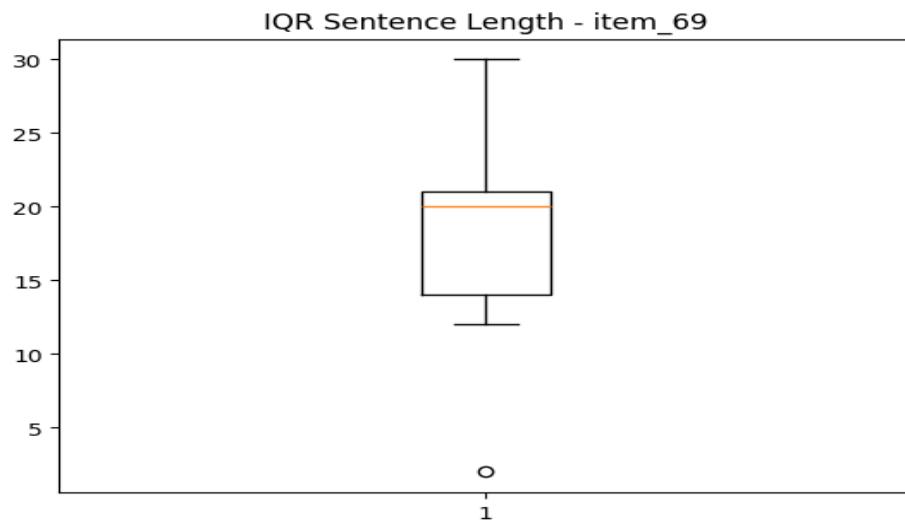
sentence_IQR_item_67.png
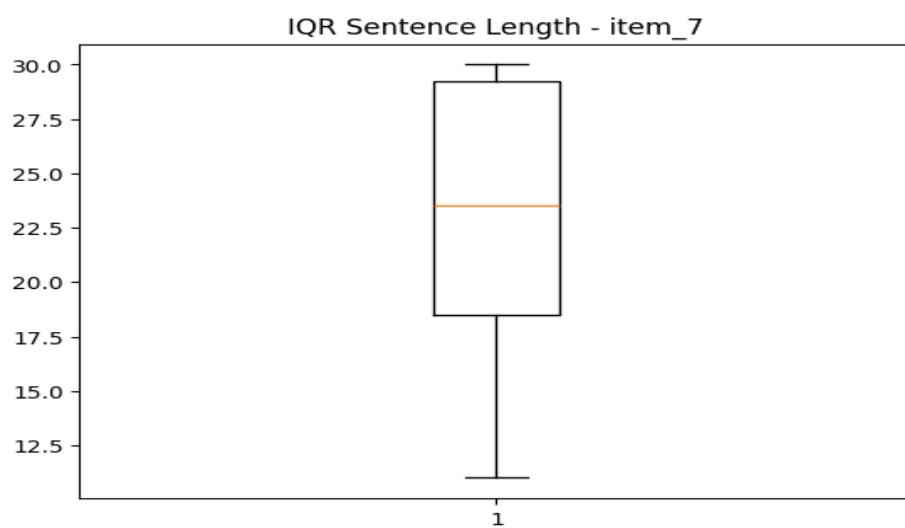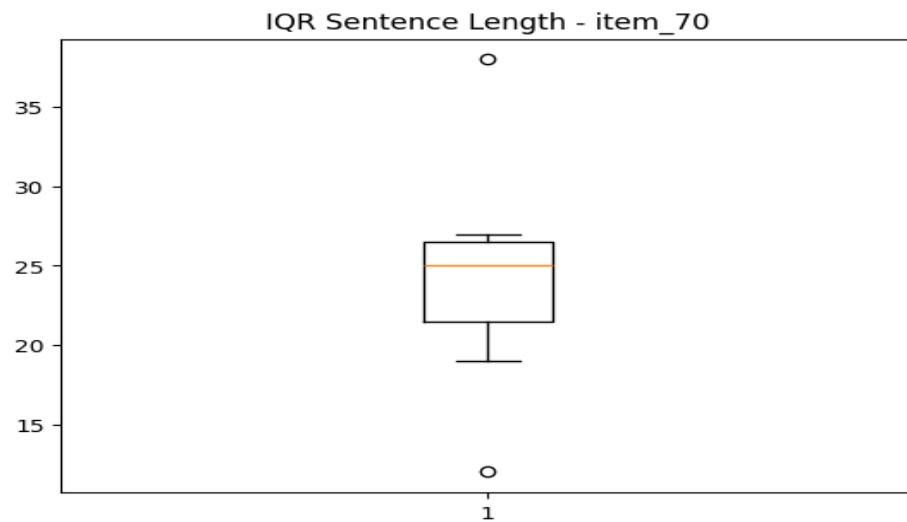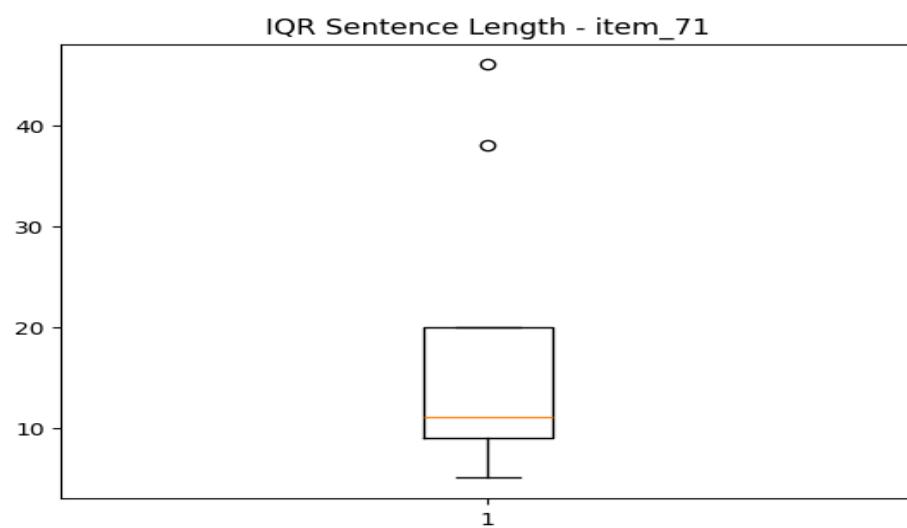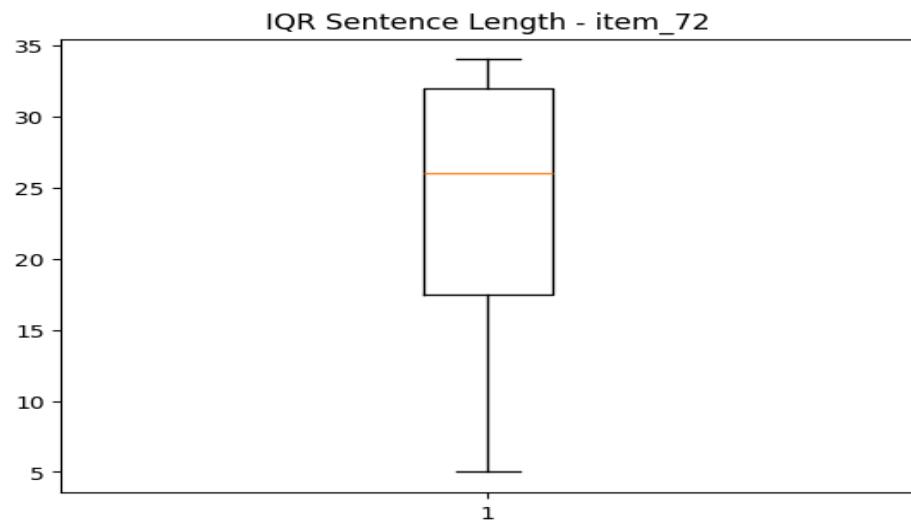
IQR Sentence Length - item_67

sentence_IQR_item_68.png

IQR Sentence Length - item_68

sentence_IQR_item_69.png

IQR Sentence Length - item_69

sentence_IQR_item_7.png

IQR Sentence Length - item_7

sentence_IQR_item_70.png

IQR Sentence Length - item_70

sentence_IQR_item_71.png

IQR Sentence Length - item_71

sentence_IQR_item_72.png

IQR Sentence Length - item_72

sentence_IQR_item_73.png

IQR Sentence Length - item_73

sentence_IQR_item_74.png

IQR Sentence Length - item_74

sentence_IQR_item_75.png

IQR Sentence Length - item_75

sentence_IQR_item_76.png

IQR Sentence Length - item_76

sentence_IQR_item_77.png

IQR Sentence Length - item_77

sentence_IQR_item_78.png

IQR Sentence Length - item_78

sentence_IQR_item_79.png

IQR Sentence Length - item_79
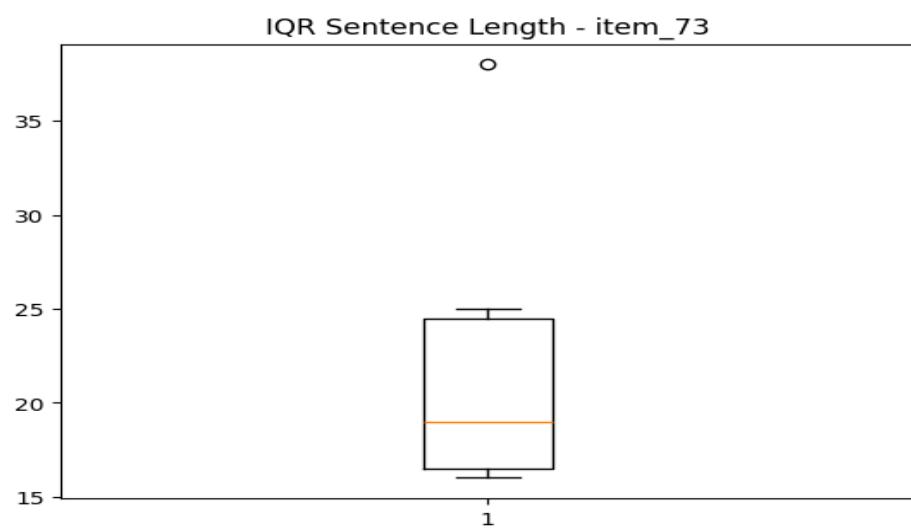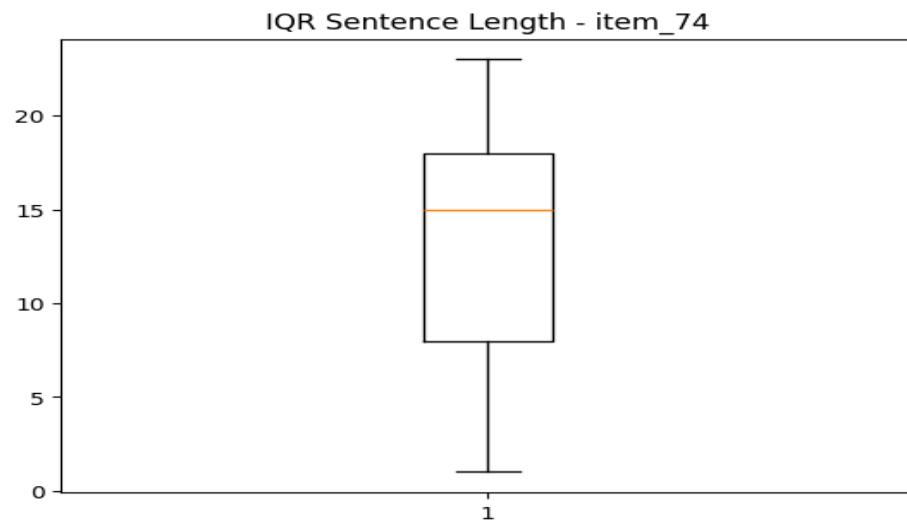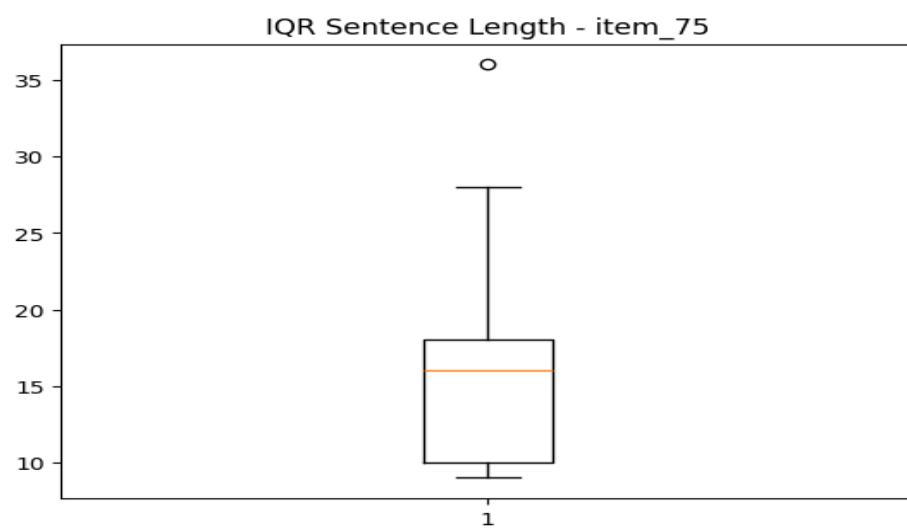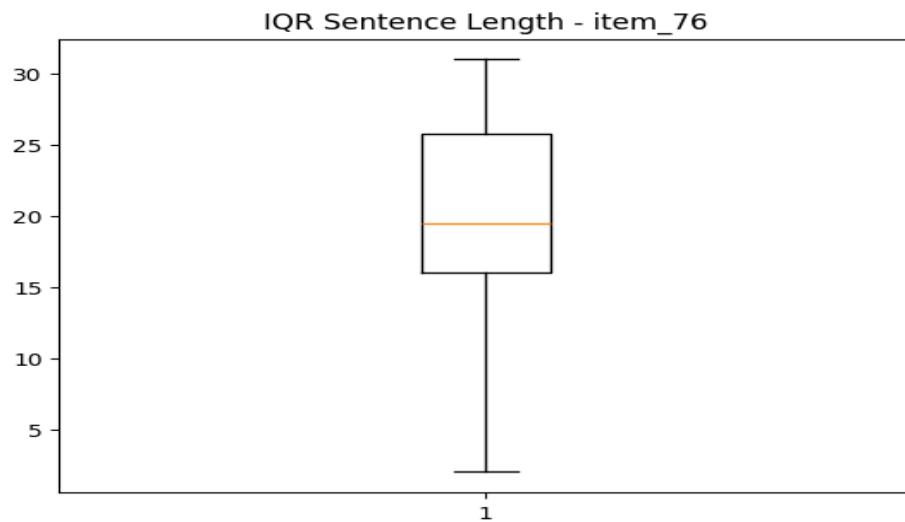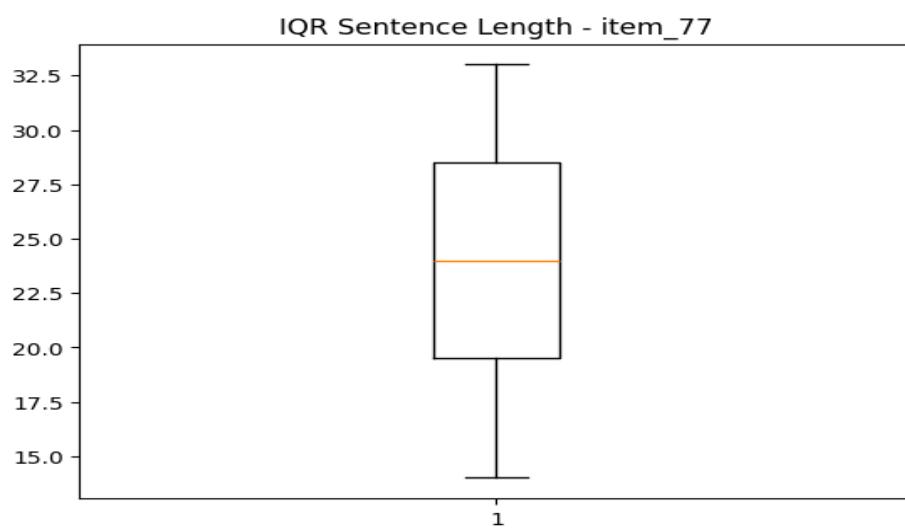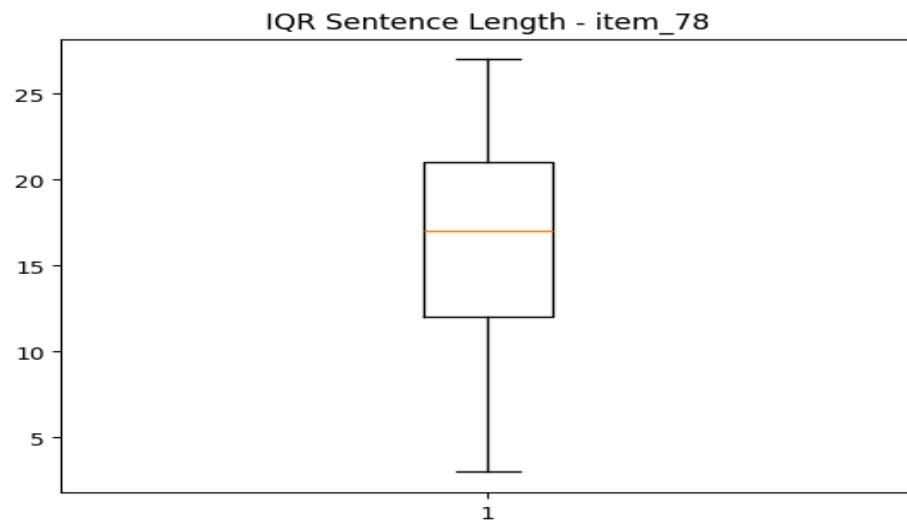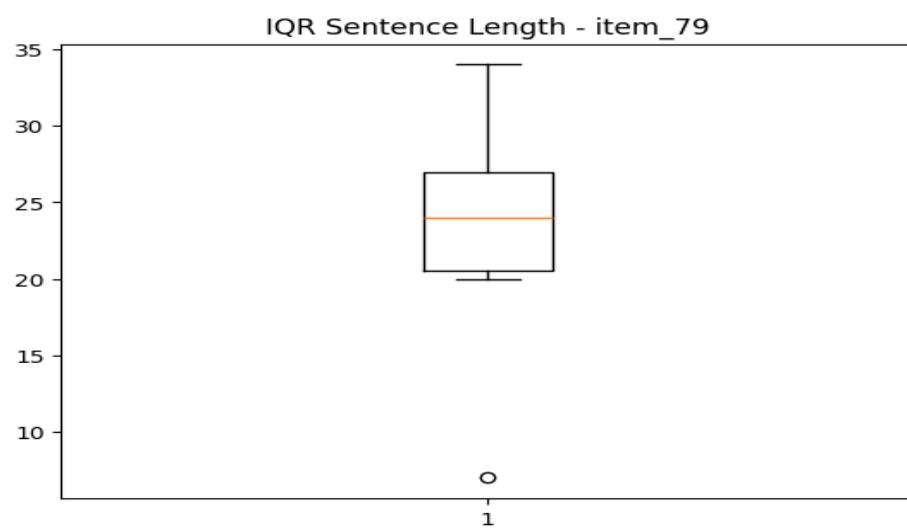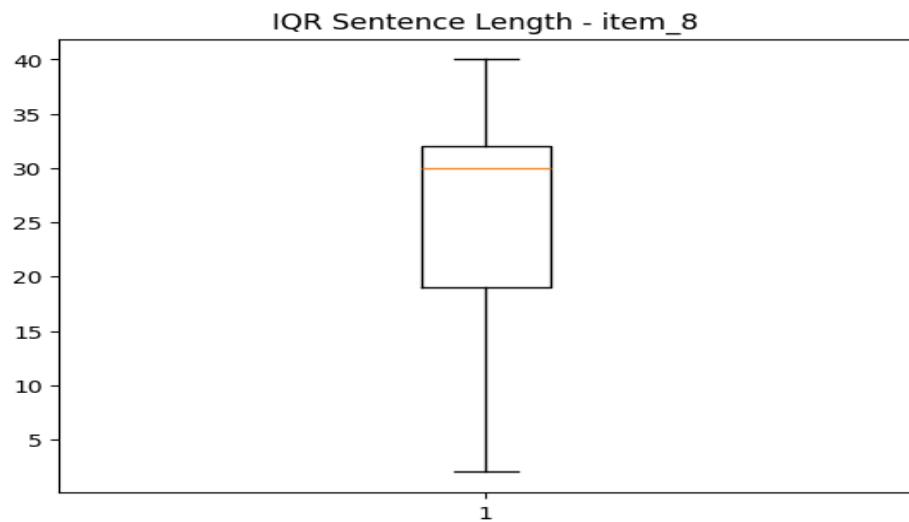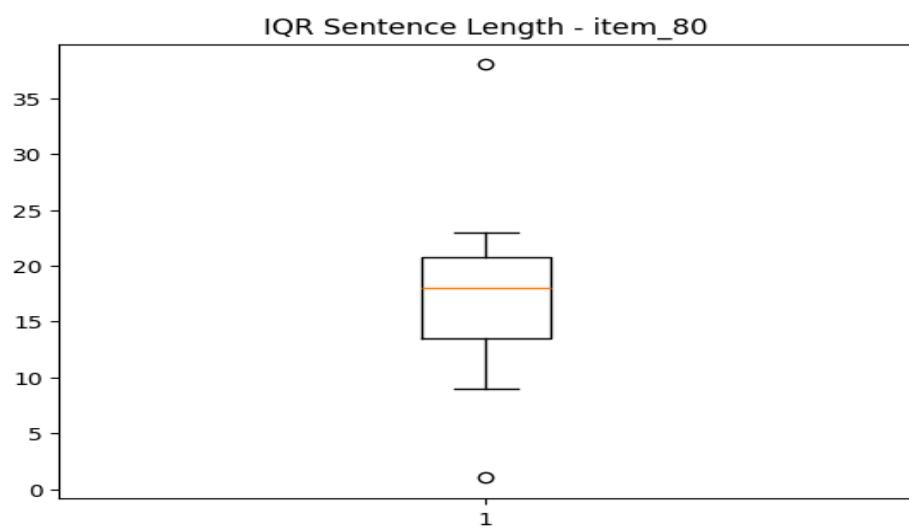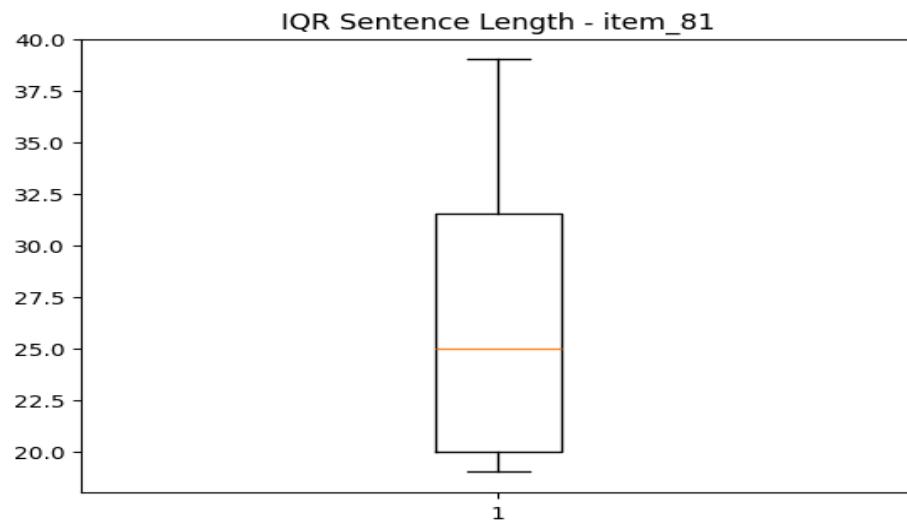
sentence_IQR_item_8.png
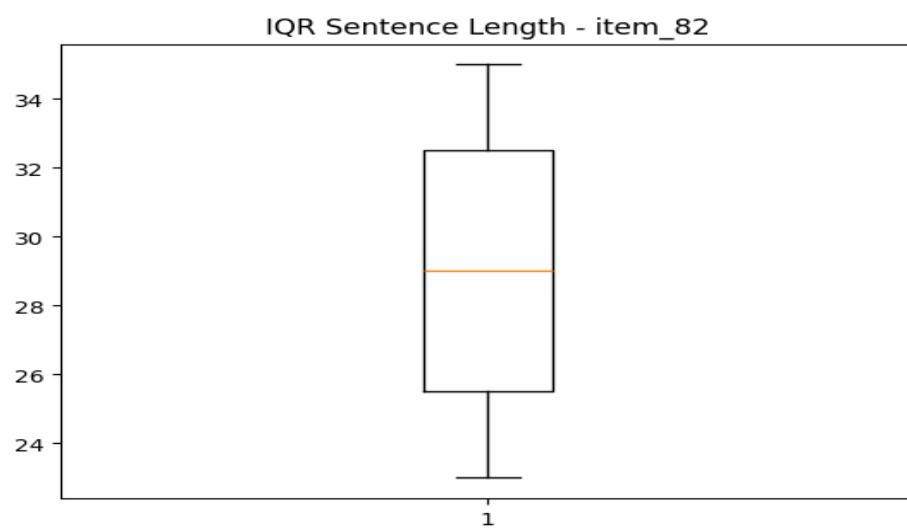

sentence_IQR_item_80.png
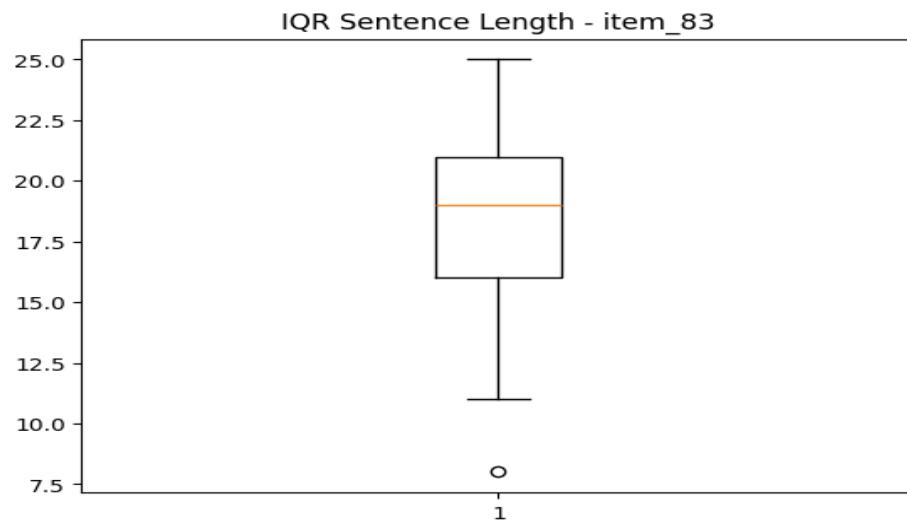
sentence_IQR_item_81.png
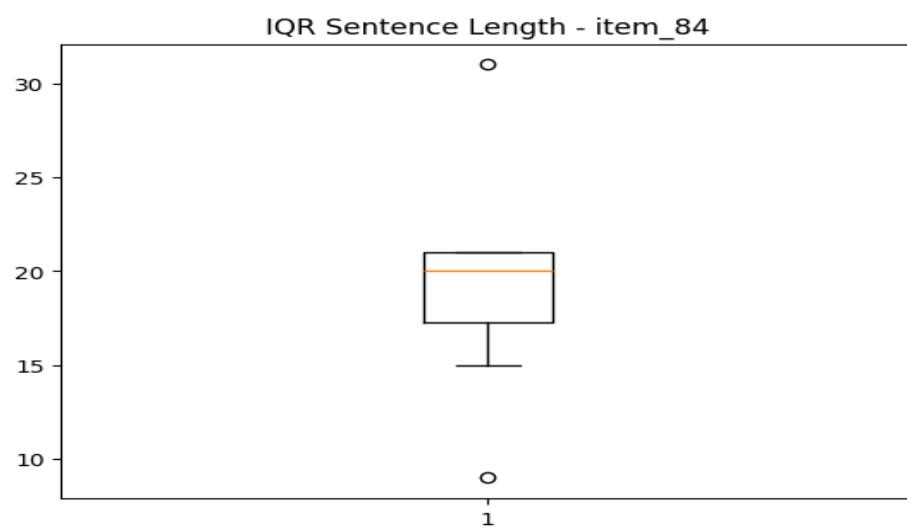

IQR Sentence Length - item_81

sentence_IQR_item_82.png

IQR Sentence Length - item_82

sentence_IQR_item_83.png

IQR Sentence Length - item_83

sentence_IQR_item_84.png

IQR Sentence Length - item_84
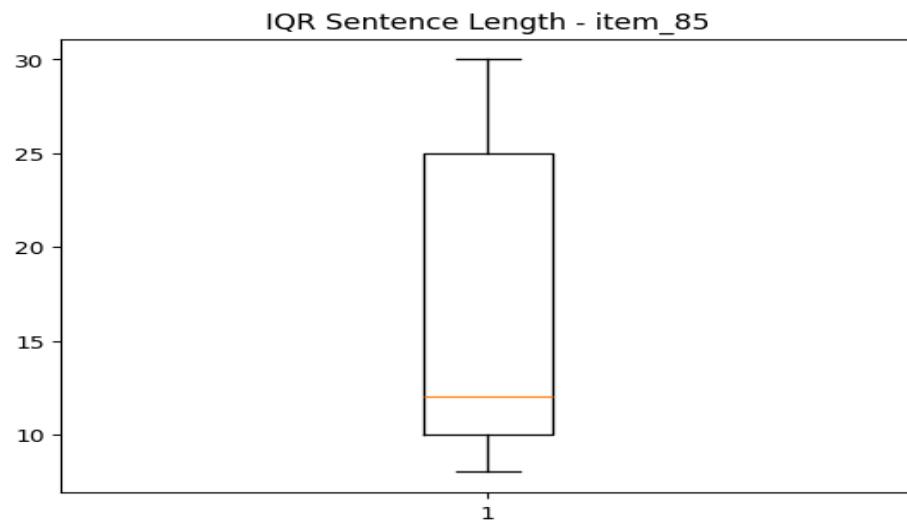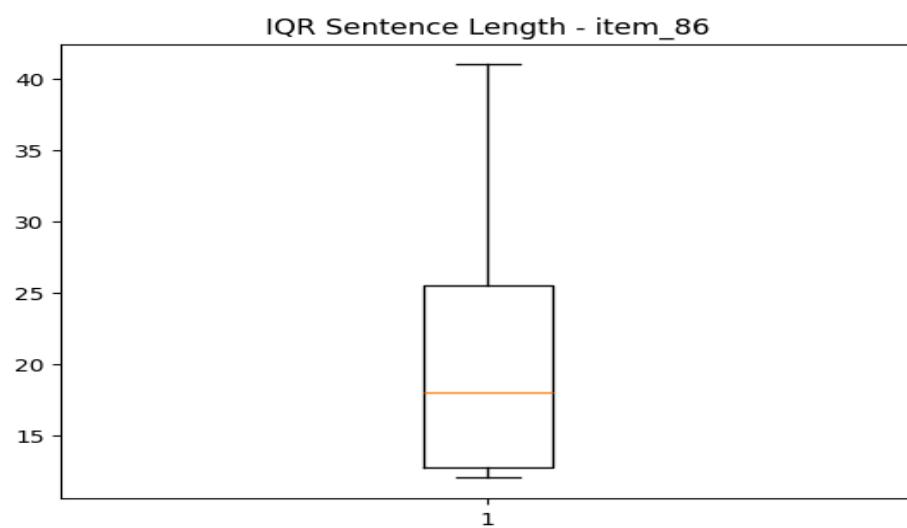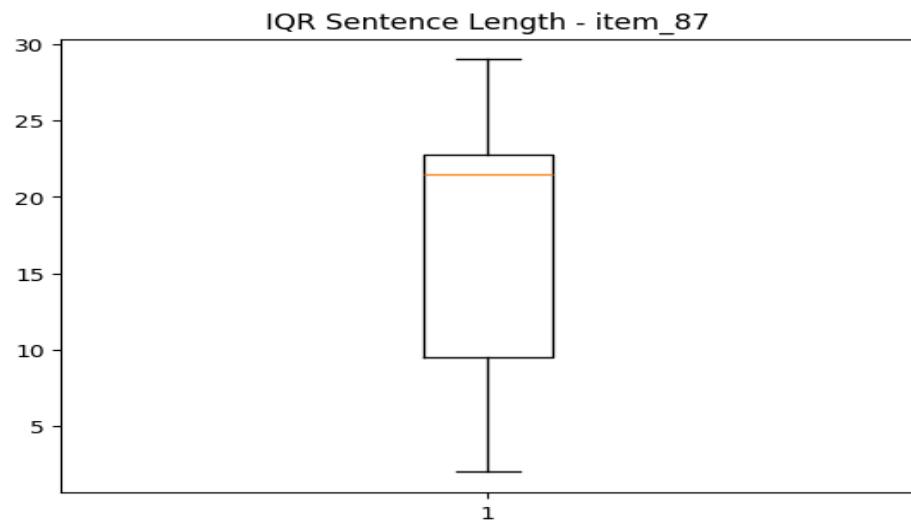
sentence_IQR_item_85.png
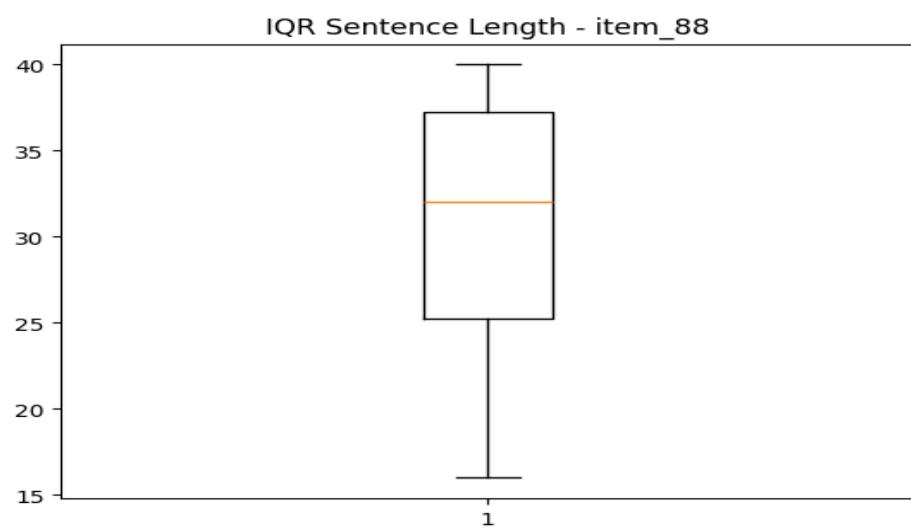
sentence_IQR_item_86.png

sentence_IQR_item_87.png


IQR Sentence Length - item_87

sentence_IQR_item_88.png

IQR Sentence Length - item_88
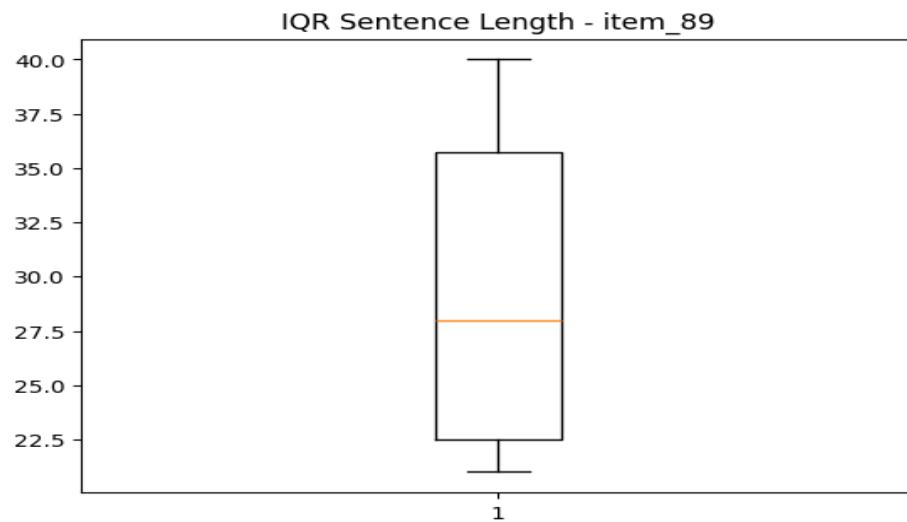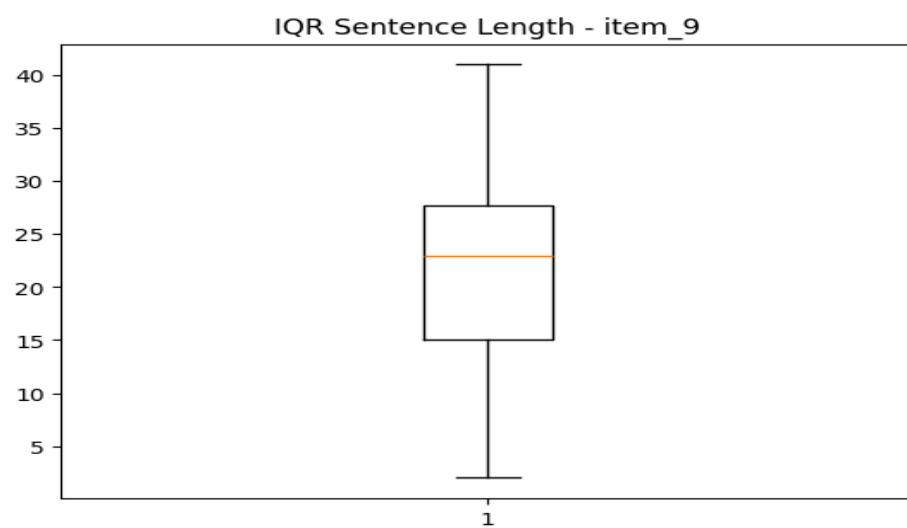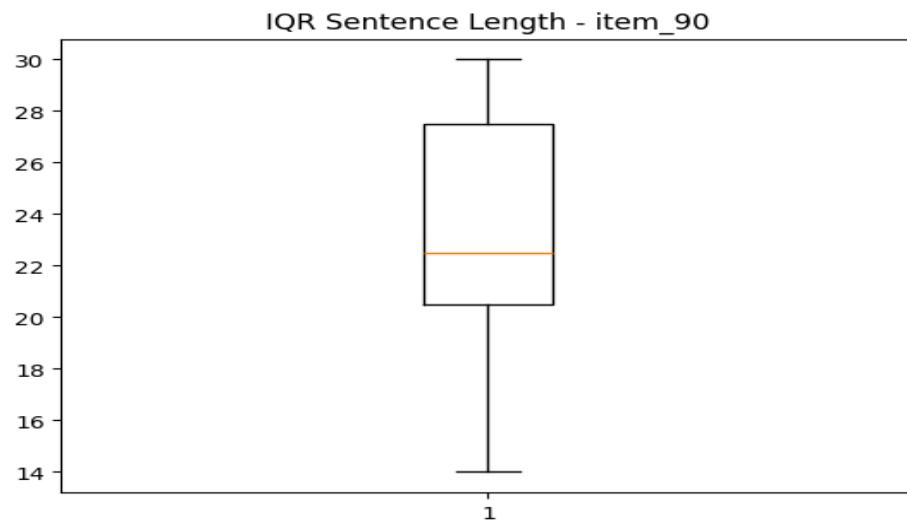
sentence_IQR_item_89.png
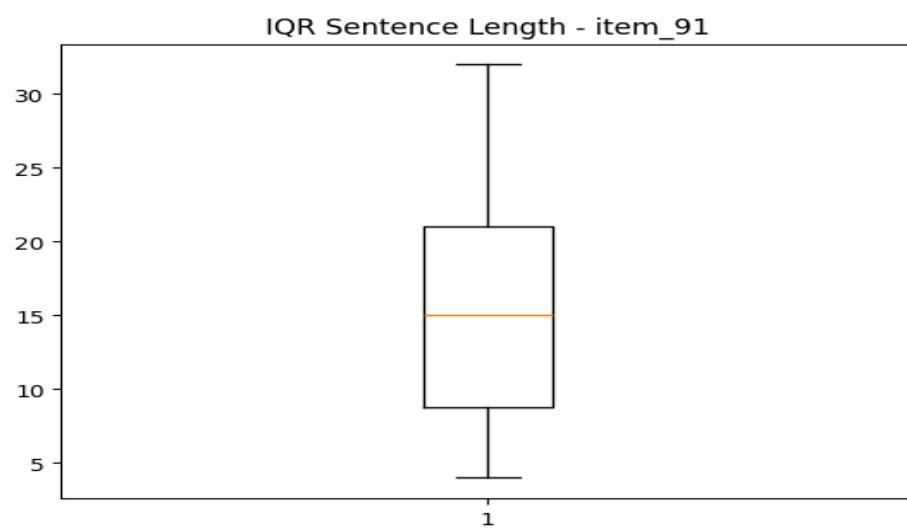
sentence_IQR_item_9.png
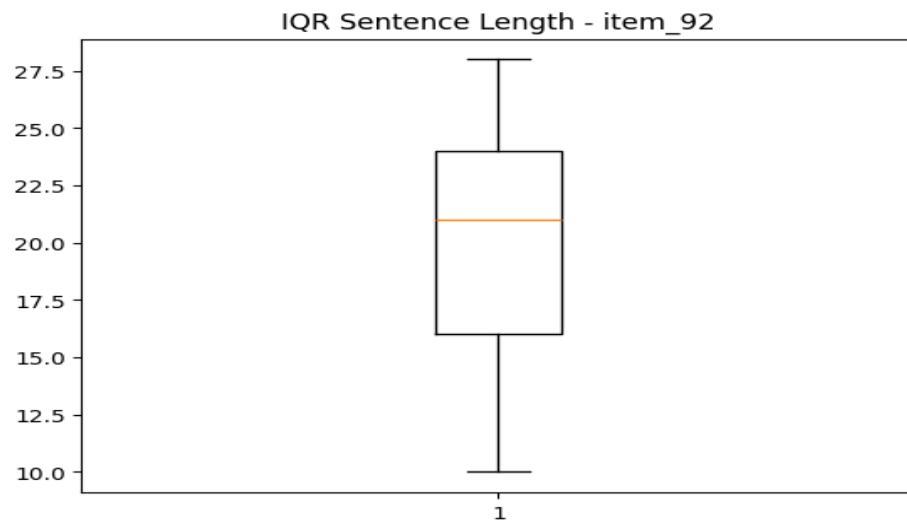
sentence_IQR_item_90.png
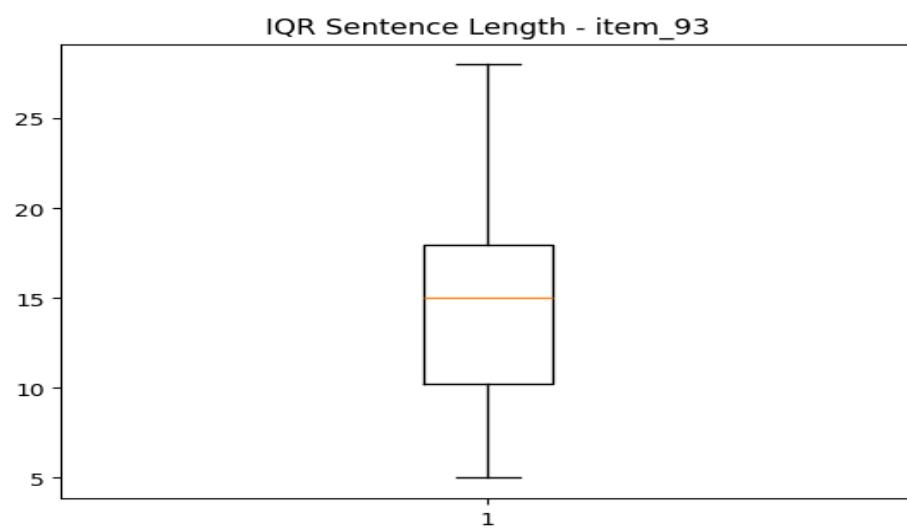

IQR Sentence Length - item_90

sentence_IQR_item_91.png

IQR Sentence Length - item_91

sentence_IQR_item_92.png

IQR Sentence Length - item_92

sentence_IQR_item_93.png

IQR Sentence Length - item_93
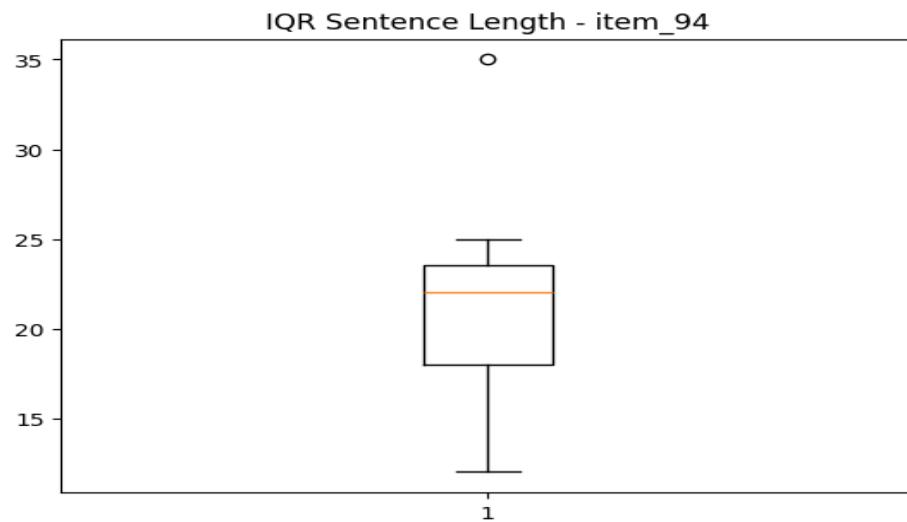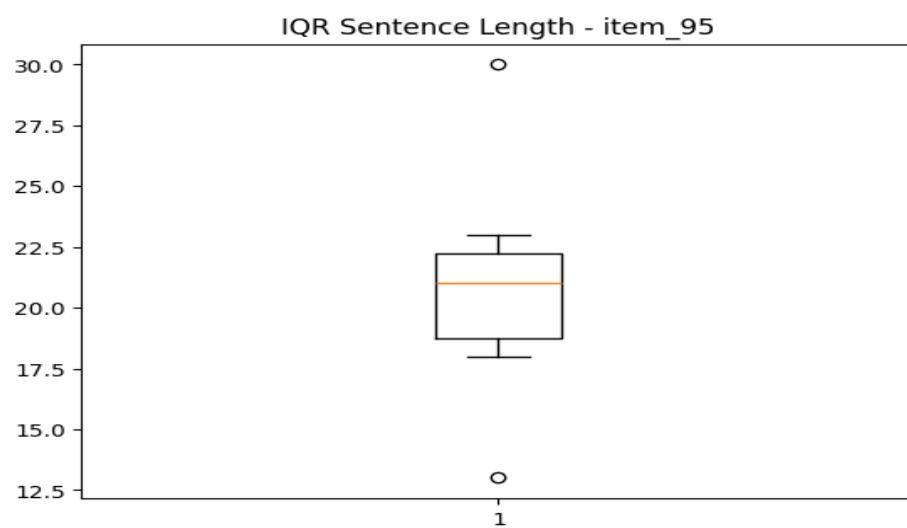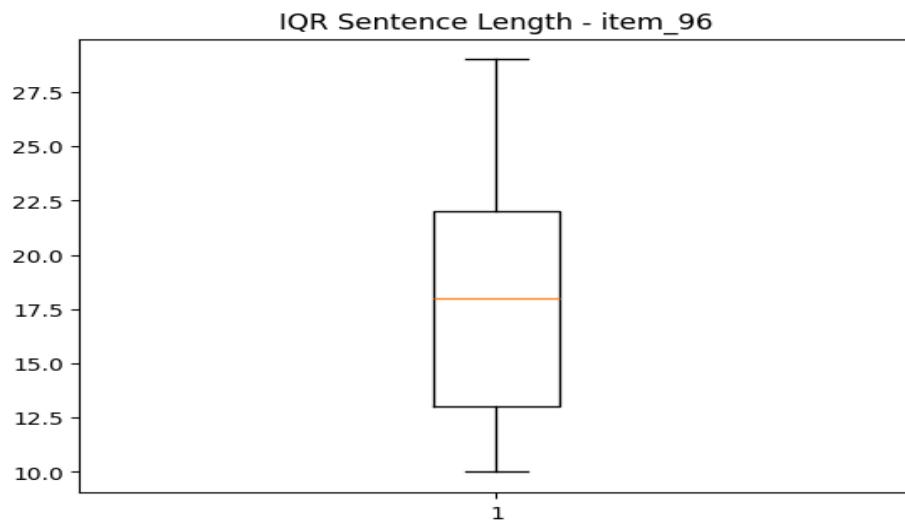
sentence_IQR_item_94.png
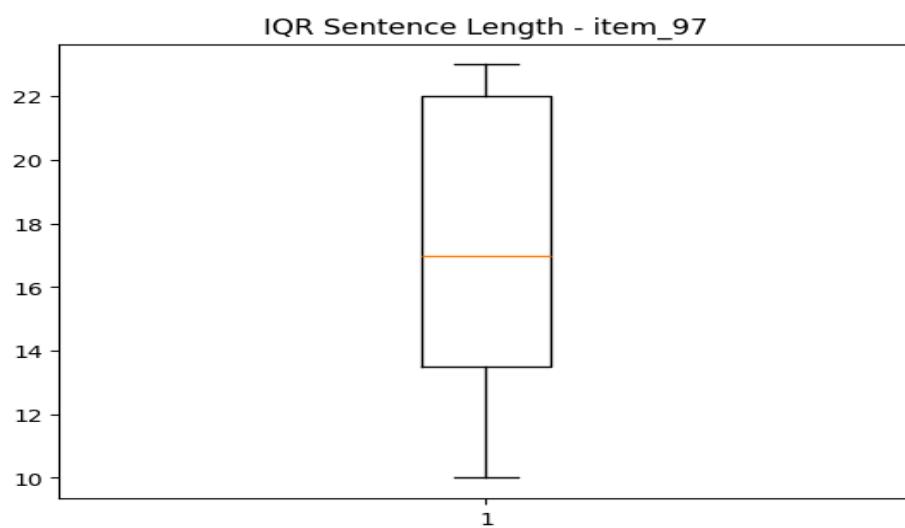
sentence_IQR_item_95.png

sentence_IQR_item_96.png


IQR Sentence Length - item_96

sentence_IQR_item_97.png

IQR Sentence Length - item_97

sentence_IQR_item_98.png

IQR Sentence Length - item_98

sentence_IQR_item_99.png

IQR Sentence Length - item_99